NASA Conference Publication 2191

Variations of the Solar Constant

S. Sofia, *Editor*
Goddard Space Flight Center
Greenbelt, Maryland

Proceedings of a workshop held at
Goddard Space Flight Center
Greenbelt, Maryland
November 5-7, 1980

National Aeronautics
and Space Administration

Scientific and Technical
Information Branch

1981

Preface

The total solar power reaching the Earth is probably the most critical parameter external to the atmosphere which affects our climate. The question of the constancy of the so-called solar constant, or solar irradiance, has been addressed for nearly a century. However, for a variety of atmospheric and instrumental limitations, a definitive answer has remained elusive.

The investigators who work on the problem of the solar constant come from very different scientific backgrounds and use very dissimilar techniques. However, this wealth of approaches has not been fully utilized since each group often ignores the efforts of the other groups which use different techniques.

When the organizing committee of this workshop first met in Spring 1980, we had in mind a number of developments which had brought this old problem again into focus. First of all, rocket-borne and balloon-borne instruments kept producing changing values of the solar constant. Secondly, indirect techniques had been proposed to monitor the solar constant, and they also suggested the occurrence of changes. Unfortunately, not only the magnitude, but even the sign of these changes were not very consistent with each other, suggesting that a straightforward interpretation of the results was not warranted. Finally, for the first time we had two spacecraft in orbit, Nimbus-7 and the Solar Maximum Mission, which could give simultaneous readings of the solar constant, and thus check the validity of each other's results. It was clear that all the workers in this subject had to talk to each other, and so the Workshop on the Variations of the Solar Constant was arranged to take place at the Goddard Space Flight Center of NASA, between the 5th and the 7th of November 1980.

The following Proceedings include all the papers read at the Workshop which had not been published elsewhere. The result is a substantially complete discussion of the solar constant problem as understood at the end of 1980.

The Organizing Committee of the Workshop had the following members:

S. Sofia, Goddard Space Flight Center, Chairman
A. Arking, Goddard Space Flight Center
J. Bahcall, Princeton Institute for Advanced Study
C. Duncan, Goddard Space Flight Center
J. Geist, National Bureau of Standards
R. Hartle, Goddard Space Flight Center
J. Mitchell, National Oceanographic and Atmospheric Administration, and
D. Wentzel, University of Maryland

CONTENTS

CONTENTS (continued)

SENSITIVITY OF THE EARTH'S CLIMATE TO CHANGES IN THE SOLAR CONSTANT

Gerald R. North
NASA/Goddard Space Flight Center

ABSTRACT

A brief review of climate sensitivity to solar variations is presented with special attention to simplified models. A number of uncertainties remain in our understanding of climate and these are elaborated upon. Especially vexing are possible feedbacks which might operate on long time scales and are therefore not testable directly.

* * * *

The earth's climate is determined by the brightness of the sun, the earth orbital parameters, and the materials of the earth-atmosphere system which dispose of the sunshine by absorbing, storing, transporting and reradiating it to space. The climate system is very complex but we can be reasonably assured that the sun is the primary forcing agent that drives it. Our crude attempts to observe and model climate suggest that the mean values of such variables as temperature are very sensitive to the solar output. Hence the climatologists are very interested in obtaining an accurate history of the solar constant.

Our ability to study climate has improved significantly in recent years because of advances in many different fields. Large computers allow us to simulate the geophysical fluid motions and forecast weather with tolerable accuracy for several days. Satelites and thousands of surface observers report data continuously; the computers assimilate the data and convert it to manageable forms. The paleoclimatic record is becoming legible through the ingenious use of tree ring data, ocean bottom stratigraphy, glacial ice cores, and other indirect methods. These are accompanied by advances in applied mathematics and statistics.

Important aspects of the earth's climate remain poorly understood despite the surge in research activity over the past few years. The scientific study of the large-scale climate is hampered by our inability to test our theories. Even the astrophysicist can test his theories of remote objects because so many different ones exist. On the other hand the student of terrestrial climatology has only one earth with a rather short and spotty record of its history.

The problem of formulating a theory which is "interesting" arises. With so few data points in space-time it is not difficult to formulate a multiplicity of theories which fit the data. All theories of climate contain fudge factors which can be adjusted to bring the model output into some conformity with observations of the present climate. The question, of course, is what happens if one of the externals, such as the solar constant, is changed. This class of investigations is called sensitivity studies. My purpose here is to familiarize you with the simplest types of sensitivity estimates in the context of the model hierarchy and to point out several paradoxes which we have not been able to resolve satisfactorily. In this way I hope you can get some feeling for the state of the art of climate modeling.

The key concept in climate modeling is that of model hierarchy discussed in reference 1. The number of variables or degrees of freedom delineate the rung occupied by a particular model. The problem is that the largest models, similar to weather forecast models, have $\sim 10^5$ degrees of freedom and the numerical solution of them proceeds at a rate of about one-tenth that of nature itself. Such models are very interesting since they are the closest facsimile we have to the real climate. We can control the boundary conditions, and do various experiments with the artificial climate produced. Still only crude attempts have been made to couple these models to the oceans or cryosphere. The reason for this lack of progress has been the longer time constants associated with ocean and ice dynamics. Any model simulating these variables will have to run for decades to reach equilibrium. Hence, the large models as presently formulated are useful only when these components are taken as given. Nevertheless, many interesting experiments are possible even with these constraints. The large models are extremely useful in establishing the higher frequency (few days) feedbacks in the atmospheric part of the climate system. The developmental research for giant models is still very actively pursued because a number of problems remain in the construction of the models especially with regard to the surface (turbulent) boundary layer and the inclusion of cloudiness variability.

At the other end of the hierarchy are the models with only a few variables. These models are motivated by the most crude expressions of the conservation of energy. The study of these models was brought into vogue by Budyko and Sellers in the late sixties. Cahalan, Coakley, and I (ref. 2) recently reviewed the progress in understanding these "toy" models over the last decade. That review was written for the general reader, hence the present note will be brief.

I can illustrate the concept of sensitivity with a simple zero dimensional global energy balance model:

$$A + BT + \frac{\sigma_o}{4}(1 - \alpha_p). \qquad (1)$$

The left hand side is the outgoing infrared terrestrial radiation (Watts/m^2), with T the globally averaged temperature, A and B are empirical coefficients estimated from satellite data. The solar constant is σ_0 (≈ 1380 W/m^2), α_p is the albedo averaged over the globe and weighted by the average fraction of sunlight reaching each latitude band; the factor of four comes from the ratio of sphere to disk area.

Consider a change in solar constant $\Delta\sigma_0$. If α_p is fixed, the change induced in T will be given by

$$\frac{\Delta T}{(\Delta\sigma_0/\sigma_0)} = \frac{\sigma_0}{4B}(1 - \alpha_p)$$

The sensitivity β_o is defined as this number divided by 100, i.e. the number of degrees of change for a one percent change in solar constant. If A and B are estimated for a black body radiator we obtain ($B \simeq 4.6$)

$$\beta_o \sim 0.60^{\circ}K$$
Black radiator

On the other hand if B is estimated from satellite data we find that $B \sim 2.0$W/(m^2deg), and the sensitivity increases to about twice that of the black body radiator. The reason for this doubling of the sensitivity is the so called "water vapor feedback".

$$\beta_o \sim 1.20^{\circ}K$$
water vapor

The mechanism responsible is connected with the empirical fact that when surface temperatures increase, the relative humidity tends to stay fixed while absolute humidities increase, thereby increasing the absorption of infrared radiation in the atmosphere. This increased greenhouse effect leads to an amplified response of the surface temperature--hence the term "positive feedback". We have confidence in our estimate of the magnitude of the water vapor feedback since its effect is almost instantaneous (therefore testable with very detailed models) and the physics is confined to radiative transfer (well understood). The net effect is simply a halving of B and therefore a doubling of β_o.

A more peculiar feedback mechanism is the ice-albedo effect. Suppose α_p is a function of temperature such that for T small α_p is large (large fraction of earth ice covered) and vice-versa. Fig. 1 shows a graph of the left and right hand sides of eq. (1); intersections of the graphs indicate equilibrium climate solutions. The warmest climate (~15°C), root I, corresponds to the present situation. Root III is a

completely ice-covered planet, and root II is an unstable intermediate solution of no physical significance. If the solar constant is lowered, roots I and II approach each other, finally merging and disappearing leading to only the deep freeze root, III. Fig. 2 shows the global temperature as function of the solar constant, as computed from Fig. 1. It is remarkable that such a simple model can exhibit so rich a solution structure. The literature over the last few years (see, for example, ref. 2) has revealed study after study up and down the model hierarchy all of whose members have this same multiple solution property due to the nonlinear ice-albedo feedback. Even the most complicated models, however, still parameterize the deposit of snow and ice in essentially the same way. No models ask about the availability of enough moisture to ice over the earth, for example.

Here we are led to the so-called Faint Sun Paradox discussed by many authors (for example, ref. 3 and 4): since fundamental astrophysical considerations suggest that the luminosity of the sun has increased monotonically from around 70 percent of its present value, we see from Fig. 2 that the earth should still be iced over. It is not iced over (!) and geological evidence suggests that it never was. The favorite way out of this dilemma is to speculate that in the past the atmospheric composition was very different with more greenhouse constituents preventing the ice over (for example, ref. 5). I find this argument very unconvincing since it requires the invocation of a rather improbable scenario. Unlike the astrophysical arguments leading to the increase of solar luminosity, it is not very "robust" regarding its dependence upon detailed assumptions that are untestable. I suggest that the explanation lies in the way ice distributes itself even on a very cold planet. In any case we are not ready to trust climate models as presently formulated under conditions more than infinitesmally different from those at present nor over such long time scales where totally new feedbacks such as geochemistry may come into play.

The presently accepted value of the sensitivity to solar constant changes is in the range 1.5 - 2.0°C. This figure is to be compared with the value 2.0 - 3.0°C change estimated for doubling the CO_2 content of the air. Most models suggest that the thermal response is latitude dependent, increasing toward the poles. Even small changes in the global temperature can be accompanied by large local effects due to the shift of climatological zones. The great plains region of the United States is such a sensitive zone.

Finally I would like to aquaint you with another paradox with conventional climate models (large and small). Geological evidence suggests that over the last few million years there have been numerous advances and retreats of the continental

glaciers on time scales of thousands of years. The period before that was ice-free and possibly 5-10°C warmer than now. The continents were in a different configuration in those days and presumably their positions now are more favorable for the cooler, more variable climate. Small climate models have been subjected to such changes in their surface boundaries and they do not yield global temperatures more than a degree or two different from the present. I suspect the same result applies up and down the hierarchy. (Thompson, Barron and Schneider reported this result as the First Conference on Climate Variations, San Diego, Jan. 1981.)

A supporting and probably equivalent paradox is related to the glacial advances and retreats that have occured over the last few million years during the recent cool period. These great waxings and wanings of the continental ice sheets appear to be in step with the changes in the earth's orbital elements (eccentricity, obliquity, and phase of perihelion) (ref. 6) on the time scale of 10^4 - 10^5 years. Again, simple climate models (ref. 2) fail to give the required responses by at least a factor of five.

It is almost certain that the large models will replicate this result, since the energy balance models are "tuned" to give the correct amplitude of the seasonal cycle and have the same overall sensitivity as the giant models.

These latest paradoxes suggest to me that low frequency feedbacks are playing a very significant role in amplifying climate response to energy budget perturbations. This problem will not be solved by improving the atmospheric component of giant models. The problem is likely to reside deep in the oceangeology system and may even involve geochemistry. Less hopeless possibilities involve the biosphere. These low-frequency feedbacks are very difficult to incorporate in our climate models, because we have no way to calibrate (fudge) the inevitable unknown coefficients. If a single coefficient is left to guesswork the whole answer is left uncertain.

If our models disagree with paleoclimatology by an order of magnitude, where do we stand on the other questions of current interest such as the doubling of CO_2, which is likely to occur in the next 50 years? I wish to emphasize that the "unknown" low frequency feedback may operate on time scales as short as decades since our closest test is the seasonal cycle.

I have tried in the foregoing to assess the current level of uncertainty in climate modeling. It is clear that we have lots of work to do to bring the data and the models together. It would be especially useful if we could ask nature to change the solar constant at different frequencies for us so we could measure the response. Remember that the system is very noisy so that we need lots of cycles (at each frequency). A similar

"natural" experiment can be done with volcanoes and the dust they leave in the stratosphere. Again we need lots of them with accurate estimates of their optical effects. My guess is that there is no substitute for monitoring and waiting.

REFERENCES

1. Schneider, S. H. and R. Dickinson, 1974: Climate modeling. Rev. Geophys. and Space Phys., 12, 447-493.

2. North, G. R., R. F. Cahalan and J. A. Coakley, Jr., 1981: Energy balance climate models. Rev. Geophys. and Space Phys., 19, 91-121.

3. Newman, M. J., and R. T. Rood, 1977: Implications of solar evolution for the earth's early atmosphere. Science, 198, 1035-1037.

4. Newkirk, G., 1980: Solar variability on time scales of 10 years to 10 years to 10 years. Proc. Conf. Ancient Sun, Edited by R. O. Pepin, J. A. Eddy and R. Merrill.

5. Owen, T., R. D. Cess and V. Ramanathan, 1979: Enhanced CO greenhouse to compensate for reduced solar luminosity on early earth, Nature, 277, 640-642.

6. Hays, J. D., J. Imbrie, and N. J. Shackleton, 1976: Variations in the Earth's orbit: pacemaker of the ice ages. Science, 194, 1121-1132.

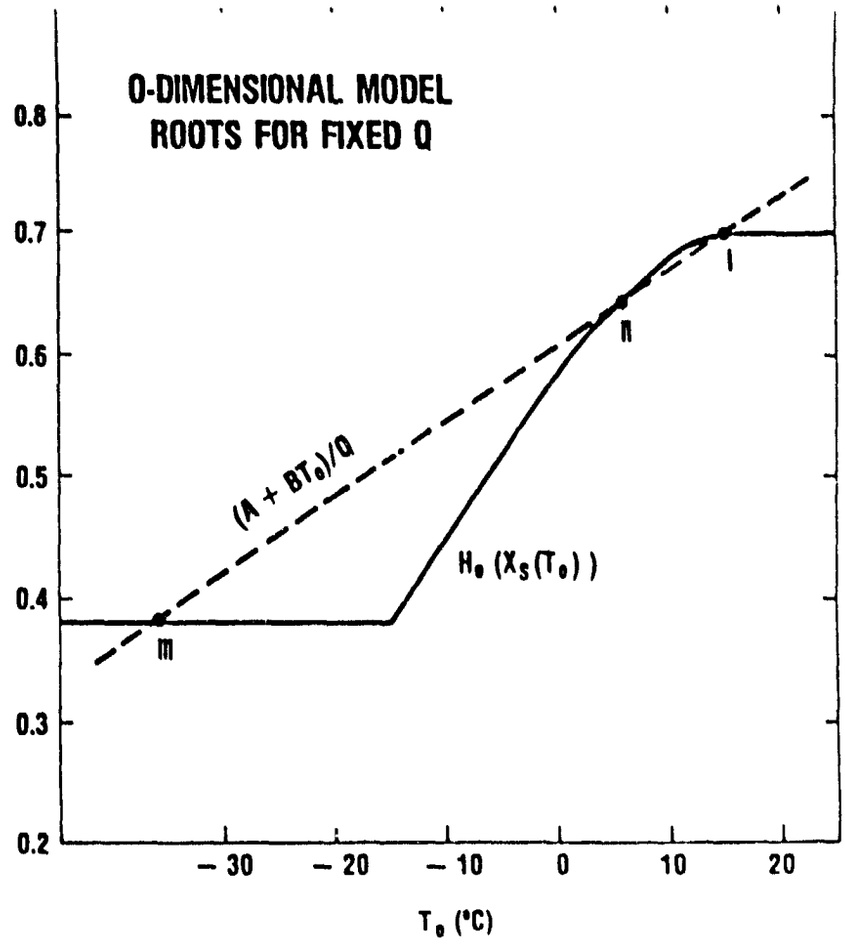


Figure 1. Fraction of solar radiation absorbed (solid curve), depicted here as H_0 ($x_s(T_0)$) but in the present paper (1-αp). The dashed curve shows the left-hand side of Equation 1 with $Q = \sigma_0/4$. The three roots are described in the text. (from Ref. 2)

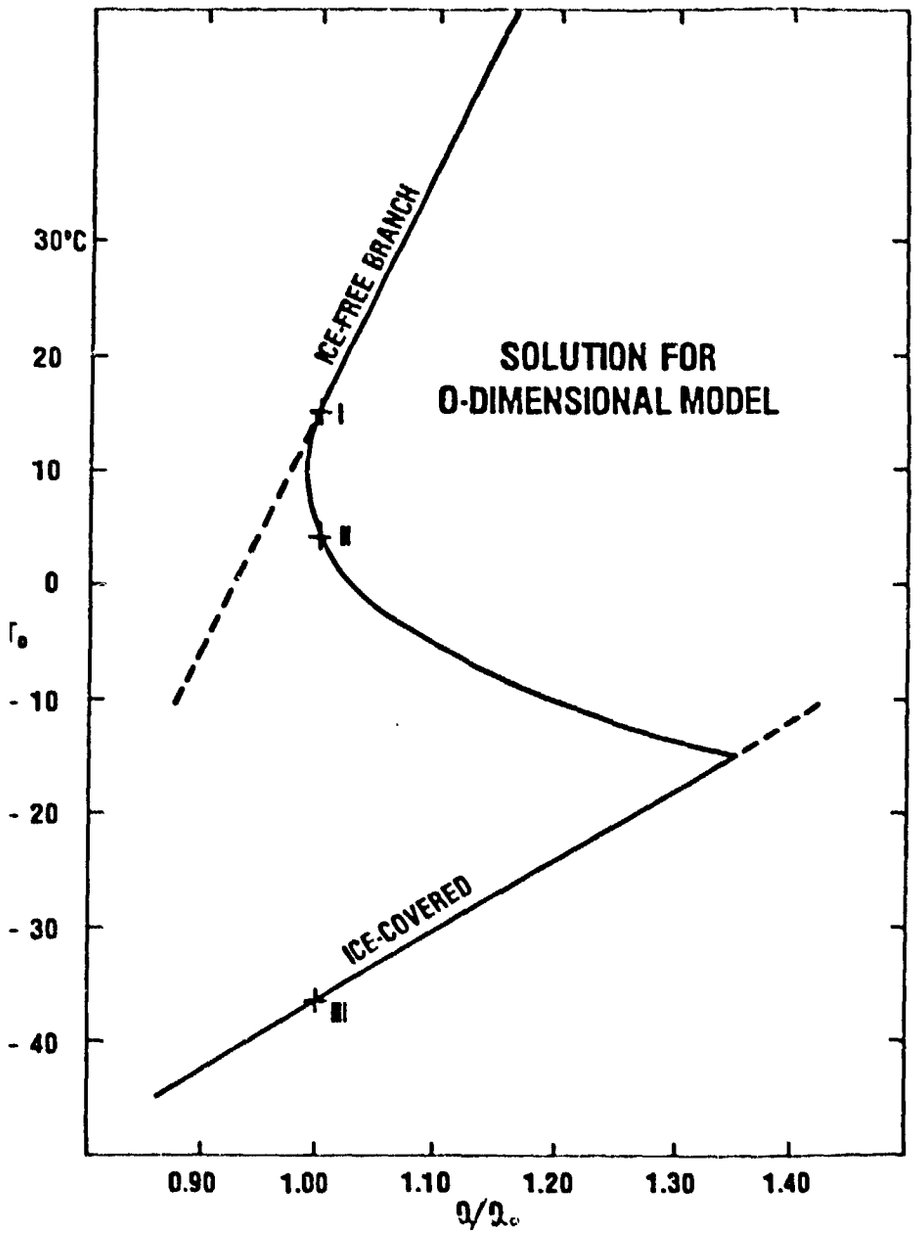


Figure 2. Steady state temperatures corresponding to solutions in Equation 1 or Figure 1 for different solar input Q/Qo (Qo is the present Q). The roots, I, II, III are the same as in Figure 1.

REVIEW OF GROUND-BASED MEASUREMENTS

Ronald J. Angione

Astronomy Department, San Diego State University

ABSTRACT

Early measurements of the solar constant made from the ground are described and discussed with particular emphasis on the Smithsonian program. A brief description is given of the monitoring program now operating at San Diego State.

HISTORICAL PROGRAMS

The solar constant is, of course, the solar irradiance at one astronomical unit integrated over all wavelengths. It is commonly expressed for example as 1368 watts/square meter, 136.8 milliwatts/square centimeter, or 1.961 calories/square centimeter/minute.

Historically the measurement of the solar constant begins in 1837 (see ref. 1). Pouillet constructed a pyrheliometer consisting of a blackened copper container, filled with water, into which a thermometer was inserted. After first determining the temperature in the shade, he placed his instrument in the sun and obtained the rate of change of temperature per minute due to the incident solar energy. Remembering that the solar constant is cal.-cm^{-2}-min^{-1}, from the known heat capacity of the water and copper he got the cal.-min^{-1}, and from the cross-sectional area of the container he got the number of cm^2, and hence the solar constant - well, almost. He had to make a correction of 2.5% for the estimated radiation lost to reflection, and he had to make a correction for the radiation lost due to scattering and absorption by the earth's atmosphere. Bouguer in 1760 had shown that for a plane parallel atmosphere the logrithm of the observed intensity is linearly related to the secant of the zenith distance (what today, allowing for curvature of the atmosphere, we call airmass). Pouillet then corrected his data for atmospheric attenuation using a linear Bouguer plot. This is not quite correct because the Bouguer plot is linear only for monochromatic radiation. Nevertheless, the value he obtained was 1.76 cal-cm^{-2}-min^{-1}, which is rather remarkable for such a simple device.

Thirty years later the famous solar astronomer, Father Secchi, built a cylindrical device with two thermometers - one exposed to the sun and one not. He used his device to determine the temperature of the sun's surface - unsuccessfully, since he arrived at a value of over five million degrees.

I mention these historical devices because (1) they illustrate the basic principles of determining the solar constant and they are remarkably like, in principle, the devices used today, and (2) the problems of determining the amount of radiation lost to scattering from the cavity and correcting for the radiation lost due to the earth's atmosphere are still the main problems of ground-based solar constant measurement.

The next major advance was the work of Langley and Abbot at the Smithsonian building upon earlier instrumental work by Langley at Allegany Observatory (Langley built the bolometer there) and Tyndall's silver disc pyrheliometer. Parallel to the Smithsonian work was the development by Angstrom in 1896 of the compensating pyrheliometer. Angstrom's device consisted of two small, thin, blackened plates of manganin (an alloy of copper, manganese, and nickel). The temperature of each plate was measured with a thermocouple, and each plate could be heated by passing a known amount of current through it. When one plate was heated by the sun, the other was heated electrically to the same temperature. The radiant power from the sun (watts) was then equal to the measured electrical power (volts times amps). You would switch the plates and take an average. The ease of operation of this device led the Smithsonian to adopt it for field work in the mid-1930's.

The bolometer, developed by Langley, consisted of two blackened strips of platinum that formed two arms of a Wheatstone bridge. When one of the platinum strips, later inclosed in a vacuum, was exposed to solar radiation its resistance changed producing a proportional change in the measured current through the bridge. On July 7, 1881 Langley's expedition with 5000 lbs of equipment (including the bolometer) left Pennsylvania by train for San Francisco (it took them 15 days), and thence with soldiers, wagons, and mules to the summit of Mt. Whitney (14,495 ft). He derived a value of the solar constant at around 3.0 cal-cm^{-2}-min^{-1} with a range of 2.63 to 3.5 (ref. 2). Abbot later reanalyzed this data set and got 2.14, which did not agree well with Angstrom's turn-of-the-century value of 1.763. Here in this difference the Smithsonian's solar constant monitoring program began.

There have been a number of published reviews of solar constant measurements: Labs at the Workshop at Big Bear Observatory (ref. 3); Thekaekara in NASA Special Publications (ref. 4 and 5); Labs and Neckel (ref. 6), and Frohlich and Brusa at the Big Bear Workshop (ref. 7) have reviewed relatively recent measurements, most of which have not been ground-based. The scales have been reviewed by Latimer (ref. 8) and Frohlich (ref. 9) among others. The two basic devices (the Angstrom and Abbot's water flow) differed by approximately 5% with the Smithsonian scale higher. Today's "absolute scale" lies just about half way between the Angstrom and Smithsonian. Abbot's improved water flow pyrheliometer with a compensating cavity (circa (1932) led to a later refinement (in 1952), which brought the Smithsonian scale

very close to today's absolute scale. Thekaekara (ref. 10) lists twelve ground-based solar constant determinations from 1940 to 1969. The range from highest to lowest is 5% of the mean, which is worse than the earlier work.

There are two basic methods employed in measuring the solar constant. One is a total measurement from exposing a blackened cavity to the sun's total energy reaching the ground. Two is a measurement of the sun's spectral irradiance, which is then integrated with respect to wavelength. There is not time to discuss all the measurements and intercomparisons in the U.S., Davos, and elsewhere, so I shall select a few to serve as examples.

In the category of "total measurement" the Smithsonian data of over 50 years (1902 - 1960) is the most extensive. Data covering a period of over a decade was obtained by Rimmer and Allen in Australia (ref. 11) using both an Angstrom and a silver disc. They reported detecting no variations in the solar constant greater than 0.1% for the period 1927 to 1939.

Sitnik (ref. 12), Stair and Ellis (ref. 13), and Labs and Neckel (ref. 14) have all determined the solar constant from spectral irradiance measurements made from mountain top sites. In all cases comparison was made to tungsten lamps and ultimately to blackbody radiation sources. Bouguer's method was used to extrapolate out to zero airmass. Labs and Neckel used a Czerny-Turner double monochrometer on a small telescope that looked at the center of the sun and then at a standard lamp, which was at the focus of a collimating mirror. Accurate knowledge of solar limb darkening was used to transfer from the sun's center to the whole solar disc. Rocket data from Tousey (ref. 15) supplied the UV shortward of 0.33 microns and model solar atmospheres supplied the data longward of 1.25 microns. They obtained a value of 1358 watts/m^2.

SMITHSONIAN PROGRAM

I will now turn to the Smithsonian program both as an example of a "total measurement" method and as an example of a solar constant monitoring program. My colleague R. Roosen and I have spent some years analyzing this data. General information concerning the Smithsonian sites is given in Figure 1. The measurements span a period of time from 1902 to 1960. The last two columns refer only to the published data. We have retrieved much unpublished data from the Smithsonian Archives for the two primary sites Mt. Montezuma, Chile and Table Mt., California. In particular the unpublished data fills in the gap from 1930 to 1940.

Figures 2 and 3 show the observing tunnels at Mt. Montezuma and Table Mt. respectively. One can see the coelostat reflecting the sunlight into

the observing tunnel, which contains the spectrobolometer. One can also see two silver discs for measuring the solar irradiance and pyranometer for measuring the solar aureole brightness.

The silver disc, shown in Figure 4, was the "field instrument". A thermometer, parallel to the viewing axis, makes a right-angle bend and protrudes radially into a blackened silver disc. One then reads the rate of change in temperature upon exposure to the sun. The silver disc had to be calibrated against an absolute detector, and so Abbot built the Water Flow Pyrheliometer, which is shown in Figure 5. The incoming water flowed across a platinum strip, whose change in resistance allowed them to calculate the starting temperature, then it flowed around a blackened cone and receiver, and finally it flowed out across another platinum strip to measure the temperature change due to the incident sunlight. They also measured the grams per minute of water flowing through the device. The Water Flow and Water Stir pyreheliometers set the "Smithsonian Revised Scale of 1913".

The Smithsonian used the spectrobolometer, shown in Figure 6, to determine the amount of atmospheric attenuation. A prism spectrograph dispersed the solar spectrum onto a bolometer. The bolometer output was measured with a Boys galvanometer, which consisted of a mirror on a quartz fiber reflecting a light beam back and forth in proportion to the signal from the bolometer. The light beam exposed a line on a moving photographic plate, which became in essence a strip chart recorder.

They typically recorded five tracings of the solar spectral irradiance on one spectrobologram, each tracing at a different airmass. The height of each tracing was measured at 34 wavelengths, and the spectrobolometer curves were numerically integrated and corrected for instrumental transmission and UV and IR losses. Since the spectrobolometer was not an absolute device, this integrated area was normalized to the total surface irradiance measured by the silver disc (later they used a Compensating Angstrom detector). Then point by point, using Bouguer's law, they extrapolated the normalized curve to zero airmass, integrated it again, added zero-airmass UV and IR corrections, and obtained the solar constant. Further discussion can be found in ref. 16.

A ground-based determination of the solar constant is only as good as the correction for the atmospheric effects on the measurement. These atmospheric effects, derived from the Smithsonian data, are shown in the following three figures. Figure 7 shows the optical depth at 4 of 34 wavelengths for Mt. Montezuma (ref. 17). One can clearly see the seasonal variation and the eruption of Mt. Quizopu, Chile in 1932. The total ozone can be determined from the Chappuis band (ref. 18), and Figure 8 shows the result for Table Mt. in the form of monthly means. Absorption by water vapor is one of the largest effects and its changes are shown in Figure 9 for both Table Mt. and Mt. Montezuma (ref. 19). They determined it from the depths

of three water vapor bands, calibrated by Fowle, and developed a scheme for calculating the band areas from these depts. From these figures one can clearly see that these atmospheric effects are both large and variable and must be accurately determined on each day of measurement. Rayleigh and aerosol scattering remove at one airmass typically 10 to 15% of the direct solar beam, ozone 4 to 5%, and water vapor 10 to 20%. In spite of all this, it is possible to make accurate atmospheric corrections, and the Smithsonian was successful in doing so.

Our on-going analysis of the Smithsonian data (ref. 16), funded by NSF, shows that they made the atmospheric corrections quite well. The main question is whether they detected variations in the solar constant. This has been much discussed in the past, but perhaps Figure 10 will allow the reader to make his own judgement. In this figure, which is a preliminary plot from our analysis, Roosen and I present the raw, daily solar constant values for both Table Mt. and Mt. Montezuma; there are several thousand data points. The values range from about 1.92 to 1.96 cal-cm^{-2}min^{-1}on the scale of 1913. If it is the case that the variations are due to the sun, then one should see the same variations, in phase, at both sites. A preliminary cross correlation analysis shows no solar variations greater than 0.3%.

MODERN PROGRAM

There are five ways in which we can now improve on ground-based solar constant measurements since the time of C. G. Abbot and the Smithsonian program. First, instrumentation has advanced giving us the absolute cavities of Willson, Kendall, PMO, and the ERB experiment to replace the silver disc. We have temperature-controlled silicon photodiodes to replace the vacuum bolometer. Second, we have microprocessors and automation. This allows one to obtain more than ten times the number of daily measurements on the earth's atmospheric effects than in the Smithsonian program, and a consequent increase in the accuracy of this correction. Third, there has been an increase in knowledge. We know more about these atmospheric effects, we can better measure them, and we can better correct for them. We know more about the UV and IR corrections and have more accurate solar spectra. Fourth, modern data can be reduced and analyzed with electronic computers allowing both more sophisticated and more accurate determinations of the solar constant. Fifth, we have now the ultimate calibration of ground-based measurements through intercomparison with the space measurements.

With these improvements in mind, we have begun at San Diego State's Mt. Laguna Observatory (6100 ft altitude) a long-term monitoring program on the solar constant. We intend to continue our program through at least one solar magnetic cycle. A flow diagram of our method is shown in Figure

11. The basic instruments consist of three absolute cavities (a Kendall Mk VI, a Willson ACR III, and a Willson ACR IV) and a filter wheel radiometer with a temperature-controlled (± 0.1 ^{o}C) silicon photodiode. The eleven narrow-band filters (FWHM 75 Angstroms) cover the range 3840 to 10100 Angstroms. They were carefully selected to cover the Chappuis band for ozone determination, the 9350 water vapor band, and part of the spectrum relatively free of telluric features for determining the amount of dust. For the purpose of calibration we have intercompared our instrument with both a Dobson and Glenn Shaw's instrument at Mauna Loa Observatory. The absolute cavities will be intercompared at Table Mt. with other cavities. In 94 intercomparisons between the Kendall Mk VI and the Willson ACR III, under various atmospheric conditions, the ACR III reads higher by about 4 percent, but this systematic difference is stable to 0.15 percent standard deviation.

With cavities in space, why measure the solar constant from the ground? First, the space experiments (such as SMM) may not run continuously for the next several decades. The ground-based measurements, carefully calibrated against the space measurements, can then fill in the periods of time when there are no space experiments. Second, even if the space experiments are continuous, and even if the sun should turn out to be essentially constant, then the ground-based measurements will still provide a valuable record of atmospheric aerosols, ozone, and water vapor.

REFERENCES

1. Abetti, G.: The Sun, MacMillan Co., New York, 1957.

2. Abbot, C. G., Fowle, F. E., and Aldrich, L. B.: Annals of the Astrophysical Observatory of the Smithsonian Institution, Vol. 3, 1913.

3. Labs, D.: in Proceedings of the Workshop: The Solar constant and the Earth's Atmosphere, Big Bear Obs., No. 0149, H. Zirin and J. Walter, eds., p. 157, May, 1975.

4. Thekaekara, M. P.: Survey of the Literature on the Solar Constant, NASA SP-74, 1965.

5. Thekaekara, M. P.: The Solar Constant and the Solar Spectrum Measured from a Research Aircraft, NASA TR R-351, 1970.

6. Labs, D. and Neckel, H.: Solar Physics, Vol. 19, pp. 3-15, 1971.

7. Frohlich, C. and Brusa, R. W.: Proceedings of the Workshop: The Solar Constant and the Earth's Atmosphere, Big Bear Solar Obs., No. 0149, H. Zirin and J. Walter, eds., p. 111, May, 1975.

8. Latimer, J. R.: Tellus, Vol. 25, pp. 586-592, 1973.

9. Frohlich, C.: in Total Solar Irradiance Monitoring Plan, NASA/GSFC, C. R. Laughlin and C. H. Duncan, eds., June, 1977.

10. Thekaekara, M. P.: Solar Electromagnetic Radiation, NASA SP-8005, 1971.

11. Rimmer, W. B. and Allen, C. W.: Mem. Commonwealth Obs. Mt. Stromlo, No. 11, October, 1950.

12. Sitnik, G. F.: Astr. Cirk. Izdav. Bjuro Astr. Soobsh. Kazan, No. 444, 1967.

13. Stair, R. and Ellis, H. T.: J. Appl. Meteor., Vol. 7, p. 635, 1968.

14. Labs, D. and Neckel, H.: Z. fur Astrophysik, Vol. 69, pp. 1-73, 1968.

15. Tousey, R.: Space Sci. Rev., Vol. 2, p. 3, 1963.

16. Roosen, R. G., Angione, R. J., and Klemcke, C. H.: Bull. Amer. Met. Soc., Vol. 54, p. 307, 1973. Hoyt, D. V.: Rev. Geophys. and Space Phys., Vol. 17, p. 427, 1979.

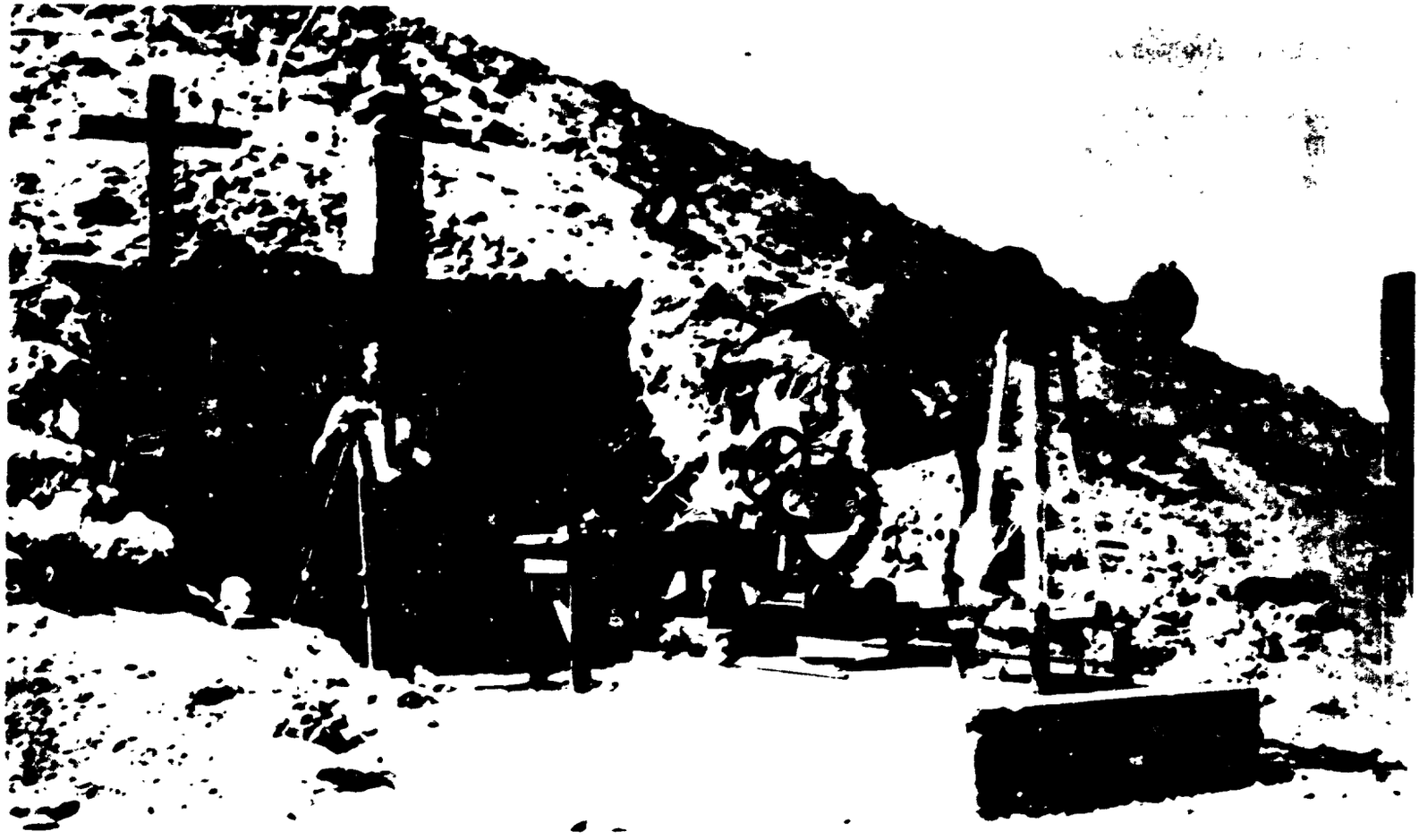
17. Roosen, R. G. and Angione, R. J.: preprint, 1981

18. Angione, R. J., Medeiros, E. J., and Roosen, R. G.: Nature, Vol. 261, p. 289, 1976.

19. Roosen, R. G. and Angione, R. J.: Pub. Ast. Soc. Pac., Vol. 89, p. 814, 1977.

SITE	LAT.	LONG.	ALT. METERS	PERIOD OF OBSERV.	DAYS OBS.
MT. MONTEZUMA, CHILE	22°40'S	68°56'W	2711	1920 - 1930	518
				1940 - 1948	46
TABLE MTN., CALIFORNIA	34°22'N	117°41'W	2286	1925 - 1930	445
				1940 - 1950	39
CALAMA, CHILE	22°28'S	68°56'W	2250	1918 - 1920	577
MT. WILSON, CALIFORNIA	34°13'N	118° 4'W	1737	1905 - 1906	121
				1908 - 1920	1124
MT. HARQUA HALA, ARIZONA	33°48'N	113°20'W	1721	1920 - 1925	225
MT. BRUKKAROS, S.W. AFRICA	25°52'S	17°48'E	1586	1926 - 1930	203
HUMP MTN., N. CAROLINA	36° 8'N	82° 0'W	1500	1917 - 1918	58
WASHINGTON, D.C.	38°53'N	77° 2'W	10	1902 - 1907	44
MT. ST. CATHERINE, EGYPT	28°31'N	33°56'E	2591	1934 - 1937	40
BURRO MTN., NEW MEXICO	32°40'N	108°33'W	2440	1940 - 1945	30
MIAMI, FLORIDA	25° N	80° W	10	1948	10

Figure 1. Basic data for the sites used in the Smithsonian solar constant program.



ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 2. The observing site at Mt. Montezuma Chile. One can see the celeostat reflecting sunlight into the observing tunnel, two silver discs, and a theodolite for obtaining the solar altitude. Courtesy of the Smithsonian Archives.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 3. Celeostat and observing tunnel at Table Mountain. Courtesy of the Smithsonian Archives.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 4. From left to right, L. B. Aldrich, C. G. Abbot, and A. Cramer. Abbot is demonstrating two silver discs and a pyranometer. Courtesy of the Smithsonian Archives.

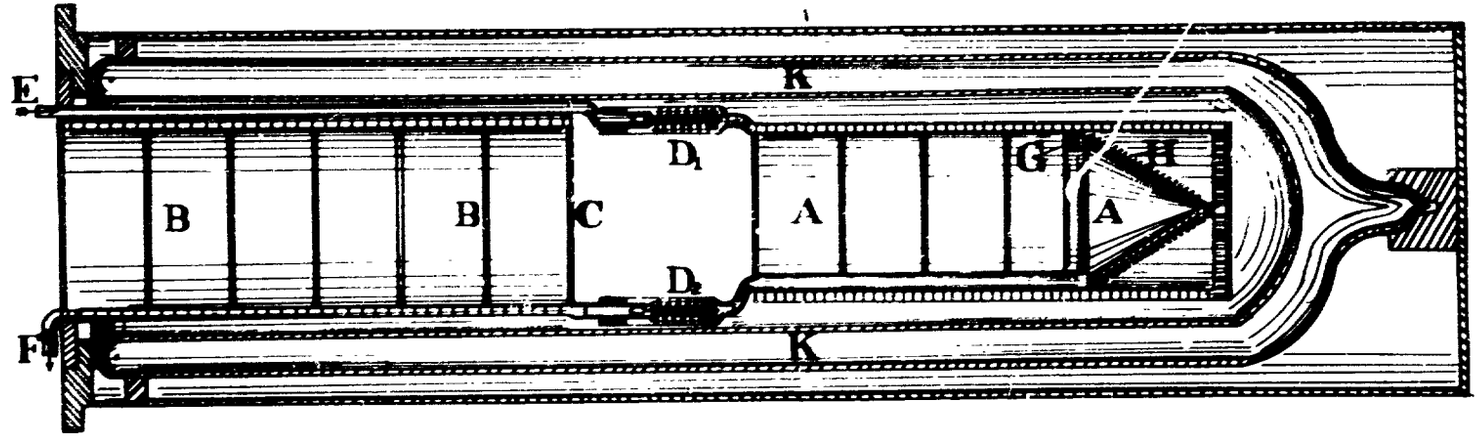
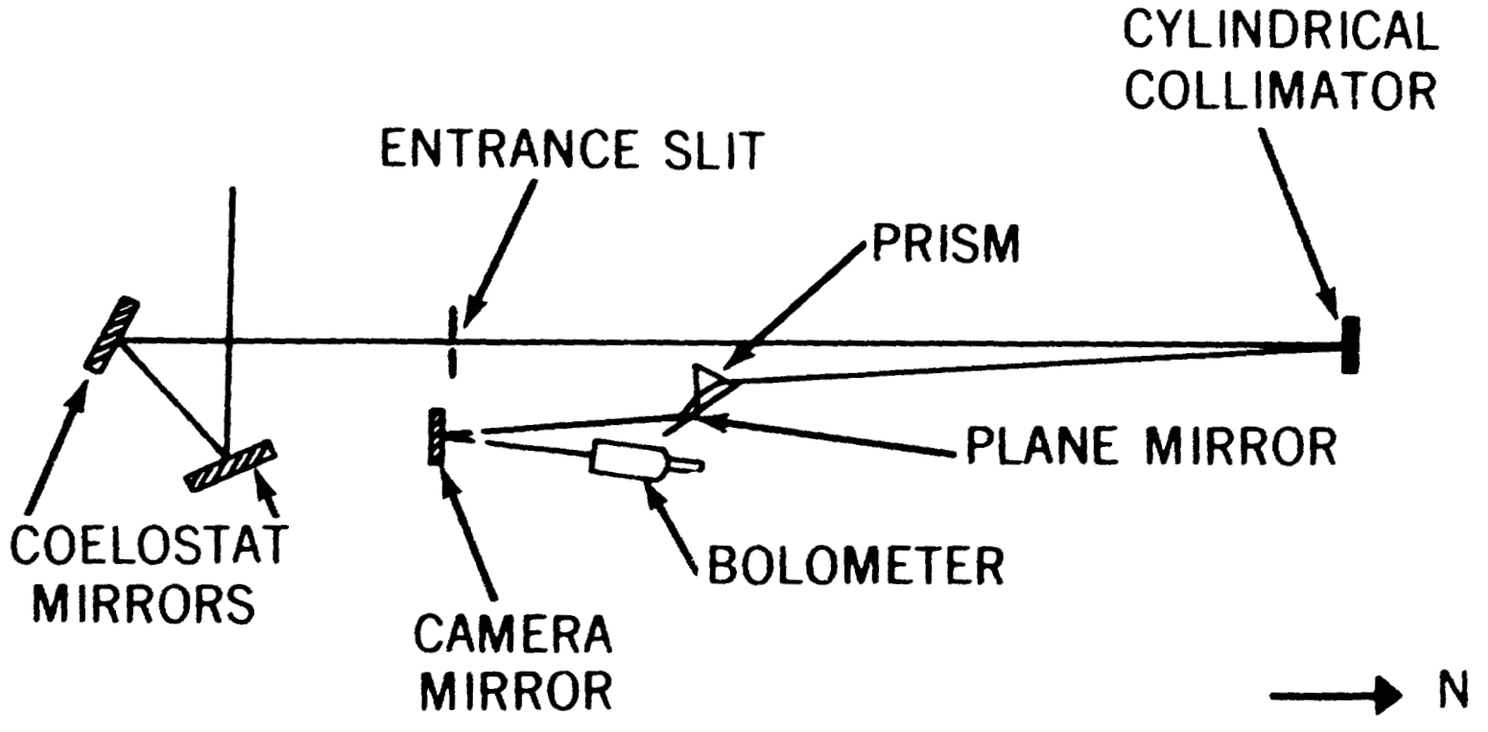


Figure 5. Schematic diagram of the Water Flow Pyrheliometer. Water enters at E and flows out at F. D_1 and D_2 are the platinum strips. Courtesy of the Smithsonian Archives.



SOLAR CONSTANT OBSERVING APPARATUS

Figure 6. Diagram of the spectrobolometer, which was used for determining the atmospheric corrections to the total solar irradiance.

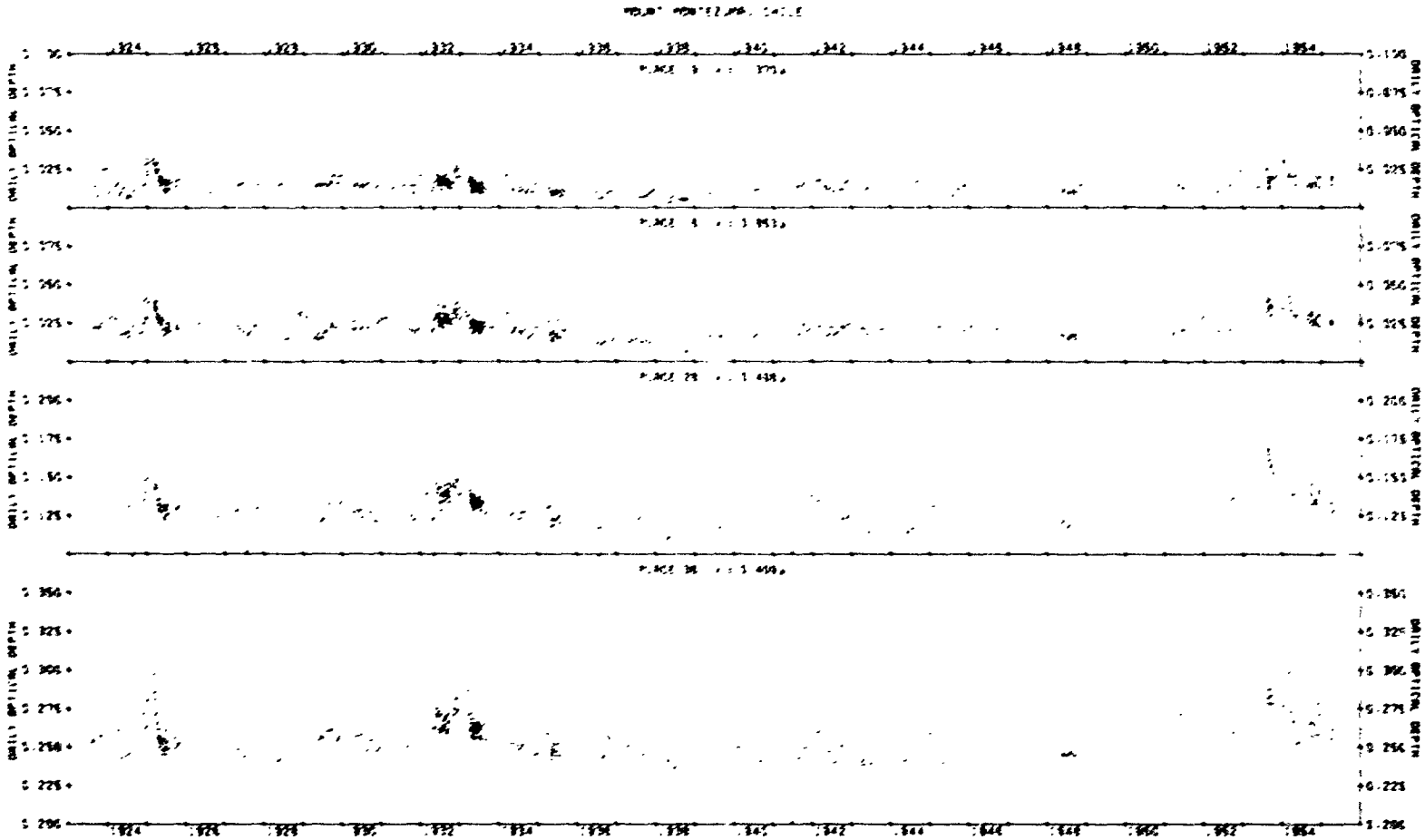


Figure 7. The optical depth at 4 of 34 wavelengths determined by the Smithsonian program at Mt. Montezuma.

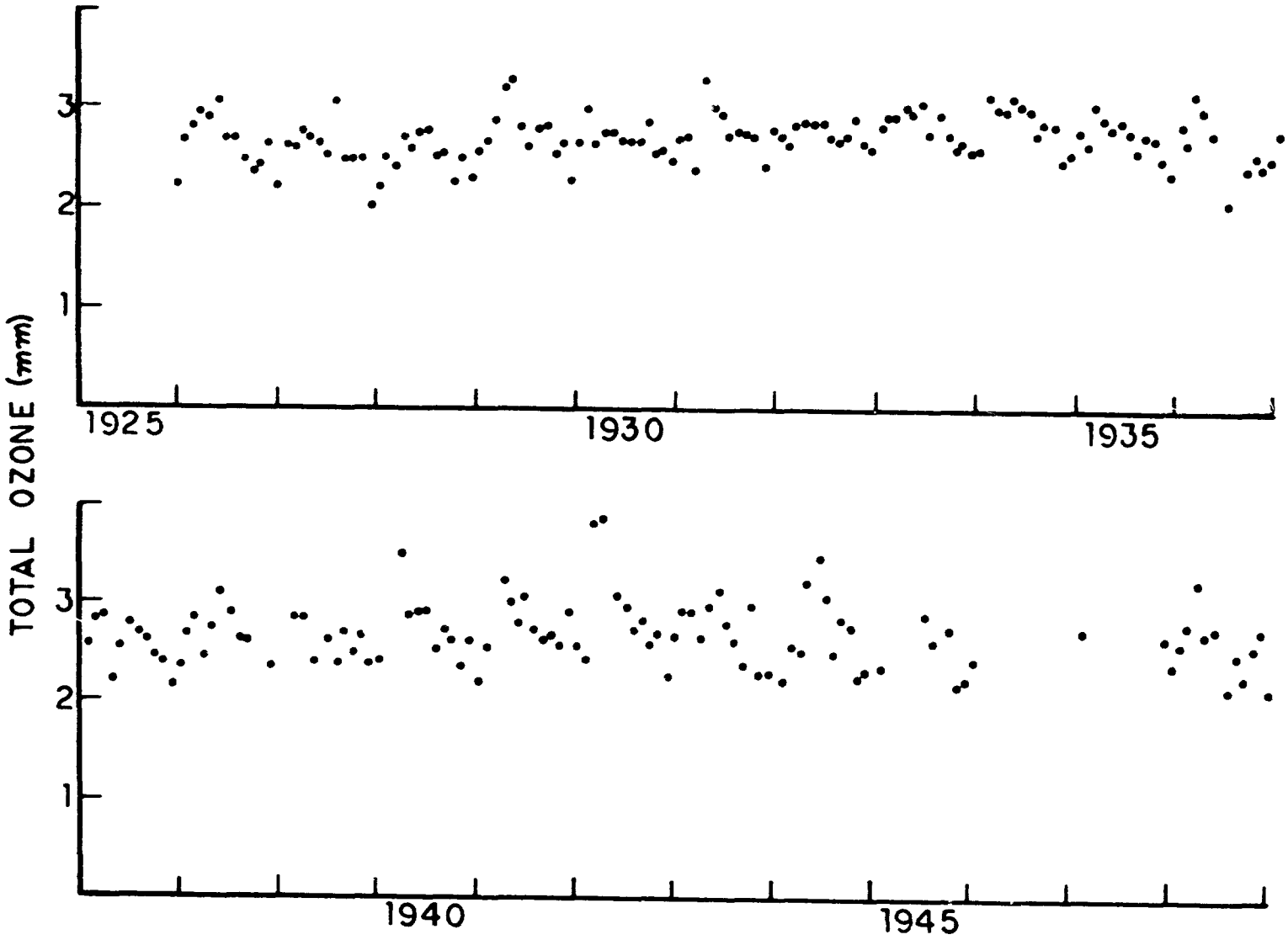


Figure 8. The monthly mean total ozone obtained from the Smithsonian data for Table Mountain. The method uses the absorption by the Chappuis band.

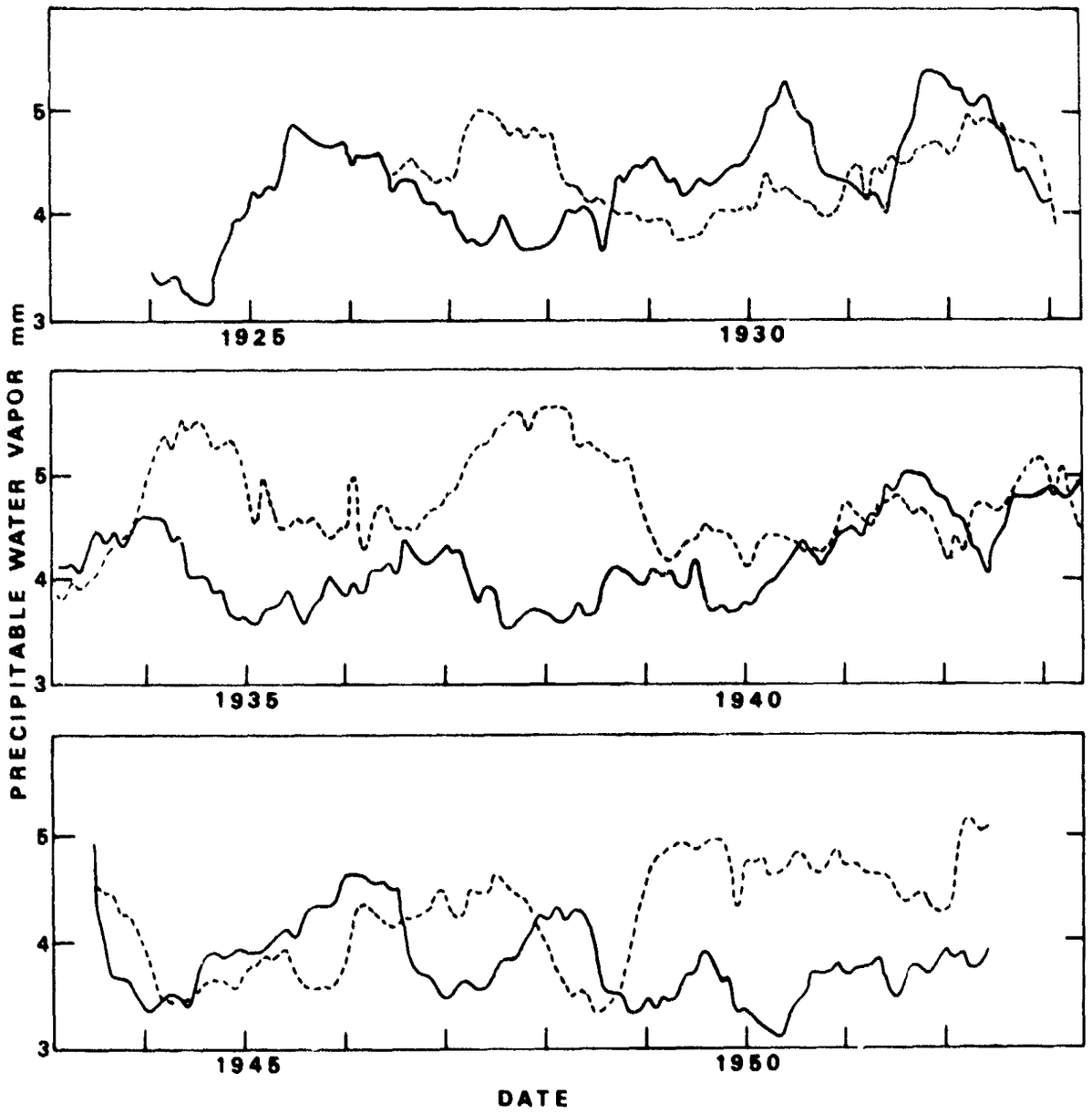


Figure 9. Twelve-month running means of precipitable water vapor at both Mt. Montezuma (solid line) and Table Mountain (dashed line).

ORIGINAL PAGE IS
OF POOR QUALITY

Figure 10. The raw, daily values of the solar constant determined by the Smithsonian program at the two primary sites.

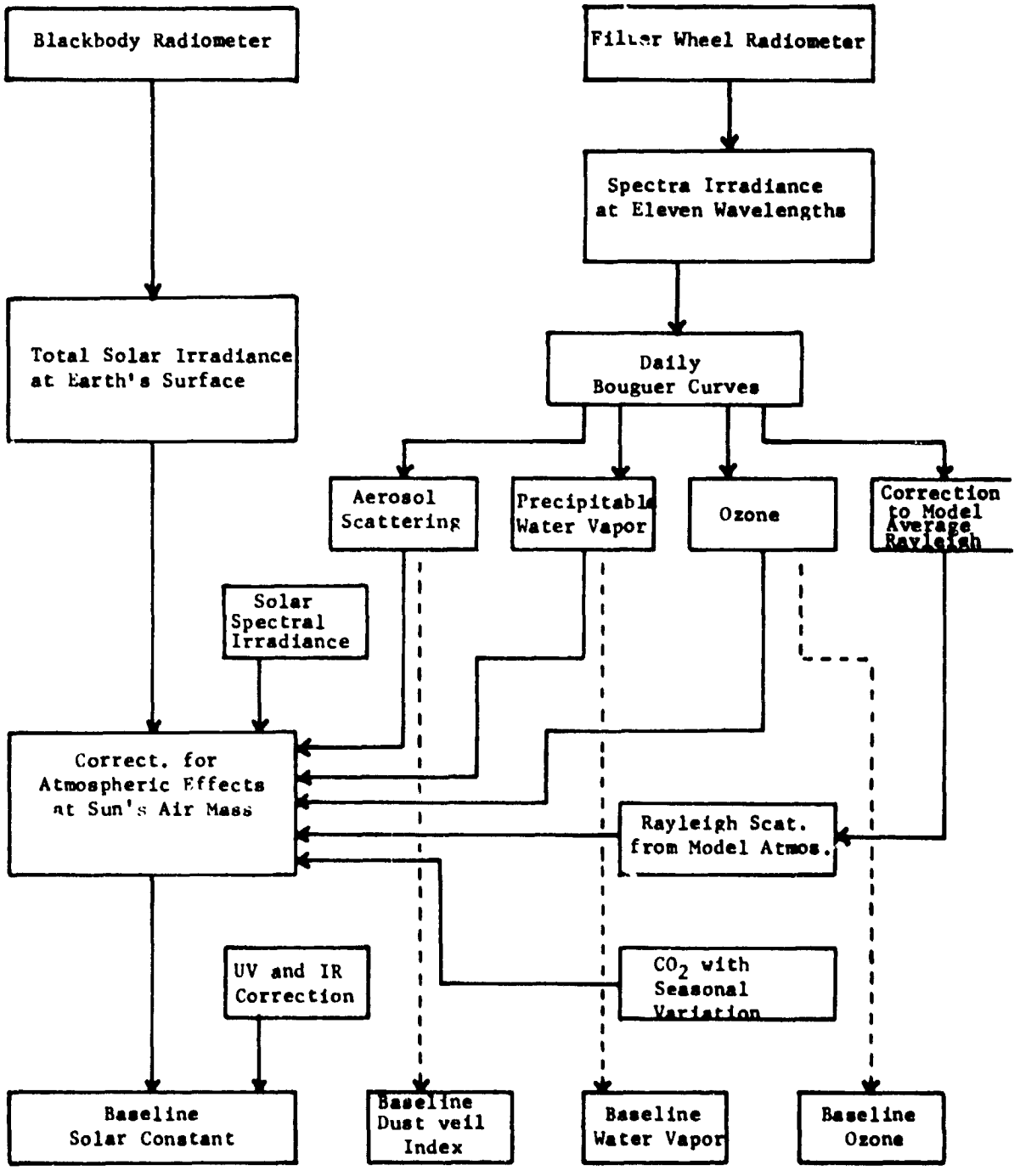


Figure 11. Flow diagram of the solar constant measurement scheme at San Diego State's Mt. Laguna Observatory.

CHANGE IN THE SOLAR CONSTANT BETWEEN 1968 and 1978*

John J. Kosters and David G. Murcray
University of Denver, Denver, Colorado, 80208

ABSTRACT

Solar irradiance measurements made from a balloon on January 27, 1978 and February 10, 1980 show a change of 0.4% over similar measurements made in 1968. This change is greater than the uncertainty of the measurement and is felt to be the result of a change in the solar constant.

INTRODUCTION

The question of the extent of variability of the solar "constant" is one that has been of interest for many years. Early attempts to assess the possible variability were limited by the large and variable atmospheric correction that had to be made to the ground-based observations used as a data base. During the 1960's several investigations were undertaken with the objective of obtaining a better measurement of the solar constant by making the observations from a high-altitude platform, i.e., aircraft or balloon (ref. 1-4). At that time we constructed a system for measuring the total solar irradiance from high altitudes using large balloons. This system was flown several times during 1967 and 1968. Once state-of-the-art measurements had been made, interest in the solar constant declined. It has been revived recently due to the current interest in climate and man's possible impact on climate. As a result of this renewed interest, we repeated measurements of the solar constant on January 27, 1978 and February 10, 1980. We used the same instrumentation for these measurements that was used in the 1967-68 measurements. The use of the same instrumentation for both series of measurements allows a direct comparison to be made of the results down to the level of repeatability of the instruments. This has been shown to be better than 0.1%. The recent measurements indicate that the solar irradiance above 30 km has increased by 0.4% over the value observed in 1967-68. In view of the measurement precision, it is felt that this change in irradiance is real and greater than can be expected due to a change in atmospheric transmission above the balloon. It is our opinion that the observed change is due to an increase in the solar constant.

INSTRUMENTATION

At the time our balloon-borne system was constructed two units were in common use for pyrheliometric measurements. These were the Eppley NIP thermopile unit and the Ångstrom-designed pyrheliometric unit. The Ångstrom unit

*This research was supported in part by the National Science Foundation under grant No. ATM 79-03877.

was designed as an absolute unit; however, it was general practice at that time to calibrate all units against secondary standards. Because the thermopile units offered a significant number of advantages from the standpoint of unattended operation, they were chosen as the units to be used in our system.

Preliminary tests with the units as supplied by Eppley indicated that, to achieve the desired precision, extensive modifications would be required. Since the units were quite susceptible to any temperature gradients, it was necessary to place them in an isothermal enclosure in order to get stable outputs. The sensitivity of the units was also temperature dependent and, although some temperature compensation circuitry was included, it was not adequate for the precision sought in the measurement.

Rather than trying to control the temperature, we calibrated the units over the range of temperature in which they would be operated during the balloon flight. Thus, the final system consisted of the pyrheliometers, a voltage reference, a low-noise amplifier system, a digital magnetic tape recording system, and a biaxial pointing system capable of orienting the two thermopile normal incidence pyrheliometers toward the sun.

The gondola housing the various units also carried the balloon control instrumentation along with a silver-cadmium 28 vdc battery supply and an antifreeze solution, with pumps for circulating the solution through the isothermal enclosure. The output of the pyrheliometers, zero reference, and the standard cell voltages were sequentially fed to a low-noise amplifier by means of geneva gear movement which operated a low noise switch. Thus all voltages were processed by the same electronics. The instrumentation is described in detail in a previous report (ref. 5).

CALIBRATION

At the time this program started, all pyrheliometric measurements were being made using the IPS 56 scale. During the late 1960's, the active cavity radiometer system was being developed at JPL and, during the last decade, these units have been used to establish a new pyrheliometric scale which is slightly more than 2% higher than the IPS 56 scale. A discussion of these scales and the relations between them is given by Frohlich (ref. 6).

Since the measurements reported here are relative, all values quoted are in IPS 56 units. The units were calibrated using an Epply Ångstrom pyrheliometer (s/n 7010). All calibrations were performed either at Echo Lake on Mt. Evans, Colorado, or at Denver. The calibration at the two locations agreed to within the precision of the measurements for similar solar intensity values. The upper range of the calibration was performed at Echo Lake since the solar irradiance levels on a good day were higher than those measured at Denver. Over the range of solar irradiance values encountered during calibration, all units agreed to within ±0.10%.

When we proposed flying the system again, we were concerned that any change which we might measure would be attributed to a shift in calibration of the units during the time they had been in storage. When the units were compared, all units agreed to within the ±0.10% using the calibration factors as determined during the earlier calibration. There was no indication of any shift in any of the units.

RESULTS

As indicated above, the units were flown four times during 1967 and 1968. All flights were successful; however, one of the pyrheliometers had a poor pressure seal and the sensitivity changed with altitude. This caused considerable difficulty in interpreting the results until after the second flight, when the problem was located and corrected.

In addition, on all flights, as the units ascend and the solar irradiance exceeds the values for which the calibrations have been performed, the two units yield slightly different results. This is due to the slight nonlinearity in the thermopile outputs. Extrapolation of the calibrations using a linear extrapolation introduces a slight error. The error is different for the two units. Thus, the irradiance values measured with P2 at float altitude are 0.4% higher than those obtained with P1.

This divergence of the values obtained with the two units - when the solar irradiance exceeds the ground calibration - can be followed throughout the ascent, and the peak difference of 0.4% occurs at float. The 1968 values of the solar irradiance as measured by the two units were P1 = 1295 w/m^2 and P2 = 1301 w/m^2.

During the 1978 flight the two units again measured slightly different values since the same nonlinear effect was present. The solar irradiance as measured by P1 was 1300 w/m^2 and by P2 was 1306 w/m^2. The same values were measured on the 1980 flight. Thus both pyrheliometers show an increase of 0.4% in the solar irradiance between 1968 and 1978, with the value remaining the same between 1978 and 1980.

DISCUSSION

The measurements of solar irradiance made with these instruments can be interpreted in three ways:

1. The change between 1968 and 1978 reflects the precision of the measurement, possibly reflecting a shift in calibration between 1968 and 1978.
2. The change reflects a real change in the solar irradiance above 30 km, the change being due to an increase in the atmospheric transparency between 1968 and 1978.

3. The change reflects a real change in the solar irradiance above 30 km with the change being due to a change in the solar constant.

As noted in the discussion above, there is no indication of a change in calibration between 1968 and 1978. All ground calibrations indicate a precision of at least 0.1%. The flights made in 1967-1968 all yielded the same solar irradiance, the two flights made since 1978 yield the same solar irradiance which is 0.4% higher than the previous results. We feel that the first interpretation is not valid, and the data indicate a real change in solar irradiance above 30 km. The question of whether this change can be due to a change in atmospheric transmission above the balloon is not an easy one to answer. We argue against it on the following basis: at balloon altitudes and high sun, the residual atmospheric absorption reduces the solar irradiance by at most 2.5%. The majority of this absorption is due to ozone with smaller amounts due to Rayleigh scattering and infrared absorption by CO_2 and H_2O. Thus, in attempting to account for the 0.4% change in the solar irradiance by a change in atmospheric transmission, a change in ozone above the balloon appears to offer the only possibility. Calculations were performed using 0.065 atm cm of ozone, a secant factor of 1.4 and the solar spectral distribution curve of Thekaekara (ref. 7). The value used for the ozone amount is typical of the amount above 30 km in the average curves given by De Luisi (ref.8).

These calculations show that 1.7% of the extraterrestrial solar radiation would be absorbed by ozone before it reached the balloon. If the amount of ozone is halved only 1.4% of the radiation would be absorbed resulting in a 0.3% increase in the solar irradiance as observed at the balloon. The fact that such a large change in ozone results in only a 0.3% change in solar irradiance tends to eliminate a change in atmospheric transparency as the cause for the observed increase in solar irradiance.

Further evidence concerning the atmospheric transmission above the balloon is available from the flight performed on February 10, 1980. For this flight the instrument complement also included a cavity radiometer system supplied by John Hickey of Eppley Laboratories. This unit yielded a solar irradiance of 1338 w/m^2 (SI units). There appears to be some disagreement among the satellite data available for the 1980 time period; the values, however, appear to fall in the range 1365 - 1370 w/m^2. This implies an atmospheric transmission-correction in the range 2.0% to 2.4%. These values are consistent with a 1.7% correction for ozone and 0.3 to 0.5% correction for residual infrared absorptions. They are difficult to explain with any lower absorption due to ozone. The 1338 value is also consistent with the values one obtains by converting our pyrheliometer to SI units. A recent comparison of our Eppley Ångstrom unit with an active cavity radiometer indicated that our values should be multiplied by a factor of 1.024 to convert them to SI units. Using this factor, a factor of 1.7% to correct for ozone absorption and 0.7% for an infrared correction (the pyrheliometers are equipped with windows which do not transmit beyond 4 μm) yields extraterrestrial irradiance in the range 1364 w/m^2 (P1) to 1369 w/m^2 (P2). Again any lower correction for the

atmospheric correction yields too low a value for the extraterrestrial flux. The higher irradiance values are best explained by a change in the solar irradiance.

REFERENCES

1. Drummond, A.J., J.R. Hickey, W.J. Scholes, and E.G. Laue, Nature 218, 259-261, 1968.

2. Willson, R.C., Journal of Geophysical Research, 76, 19, 4325-4340, 1971.

3. Murcray, D.G., J.J. Kosters, T.G.Kyle and P.R. Gast, Tellus XXI, 620-624, 1969.

4. Kondratyev, K.Y., and G.A. Nikolsky, Quarterly Journal of Royal Meteorological Society, 96, 509, 1970.

5. Kosters, J.J., P.R. Gast and D.G. Murcray, Scientific Report No. 2, AF Contract 19628-4177, 1968.

6. Frohlich, C., The Solar Output and its Variation 93-109 O.R. White, Editor, Colorado Associated University Press, 1977.

7. Thekaekara, M.P., Solar Energy, 14, 109-127, Pergamon, 1973.

8. De Luisi, J., C.L. Mateer and C.C. Porco, NOAA Technical Memorandum, ERL-86, 1980.

D4

THE VARIABILITY OF THE SOLAR OUTPUT *

Claus Fröhlich
Physikalisch-Meteorologisches Observatorium
Davos Dorf, Switzerland

ABSTRACT

A review of recent solar constant determinations and measurements of its spectral distribution is presented. For the period from 1966 to 1980 a mean value of 1367 Wm^{-2} is determined. Within the corresponding uncertainty, no significant change of both, the integral value and the spectral distribution can be detected. However, shortterm solar variations and their spectral dependence have been deduced from measurements during four hours on June 20, 1980 from 34 km altitude with amplitudes of ±500 ppm at 368 nm, of ±200 ppm at 500 nm and ±150 ppm at 778 nm. C mparison with simultaneous total irradiance data of the Solar Maximum Mission (SMM) shows a high correlation which indicates the solar origin. The power spectrum shows a weak peak at about 3.2 mHz, which corresponds to the frequency of the 5-minutes solar oscillation.

INTRODUCTION

In the past years, many speculations have been made about the possibility of longcerm changes of the solar constant, mainly for the explanation of the Earth's climate during the past 50 to 500 years (ref. 1 and 2). The response of the Earth's temperature, which is one parameter of the global climate to a change in the solar constant, is only of about 1^{o} for a one percent change (ref. 3). Since measured changes in the Earth's temperature are of the order of a few tenths of a degree, the solar constant should be monitored to better than about one or two tenths of a percent in order to establish or prove experimentally this response. Such an accuracy is now achieved by the best radiometers, developed in recent years. Unfortunately, the period, during which such accurate measurements are available is too short to establish such a relationship from measured data.

* Part of this work has been supported by the Swiss National Science Foundation, under contracts no 2.069.78 and no 2.287-0.79.

Up to now, only energetic considerations have been taken into account. Changes of the solar output are most probably wavelength dependent. This is confirmed by the detected variability of the extreme ultraviolet and the ultraviolet (ref. 4, 5 and 6). However, no evidence of any variability in the blue, the visible and the near infrared part of the spectrum has been found. This is mainly due to the few measurement periods and to the fact that spectral instruments measure much less accurately than total radiometers. Only recently, spectral measurements from stratospheric balloons have shown the wavelength dependence in the energetic bulk part of the spectrum of the solar output, at least during short period variations.

VARIABILITY OF THE SOLAR CONSTANT

Reviews of the solar constant determinations up to the late sixties (ref. 7) and for more recent measurements (ref. 8) have been written. The results are shown in table 1 and figure 1. All results have been reduced to a common radiometric reference, the World Radiometric Reference (WRR), in order to ensure comparability and temporal homogeneity of the record. Usually, the reduction is based on results of direct comparisons between flight instruments and an instrument, directly traceable to WRR. However, it has been found, that the sensitivity of radiometers can be influenced by air pressure and thus may yield different results on ground and in space. Therefore, the results of ground-based comparisons cannot always be used for the reduction of the results to WRR. This is the case for the ACR and ACRIM type instruments of Willson (ref. 9), for SMM results and for the first rocket flight in 1978, when a hard vacuum was maintained during the measurements. These two results are directly based on their individual electrical calibration as indicated in table 1 and figure 1. As the WRR is representing the

Date	Platform	Instrument	Radiometric Reference	Solar constant value (Wm^{-2})
8/1969	Balloon	ACR 111	WRR	1369
6/29/1976	Rocket	ACR 402A	ACR	1368
11/16/1978	Rocket	ACR 402A	WRR	1365
6/14/1979	Balloon	PMO6-9	WRR	1366
5/22/1980	Rocket	ACR 402A	WRR	1365
6/20/1980	Balloon	PMO6-9	WRR	1367
2-7/1980	Satellite	ACRIM A	ACRIM	1368
			mean	1367
			Standard Deviation	1.6

Table 1: Summary of revised solar constant determinations from 1969 to 1980.

SI-units within ± 0.3% and as the absolute accuracy of the ACR and ACRIM is in the order of ± 0.2%, the comparability with WRR values is at least within these limits, although the results indicate that it is probably better. During the other two rocket experiments, no vacuum was present, hence the results of the ground comparisons can be used.

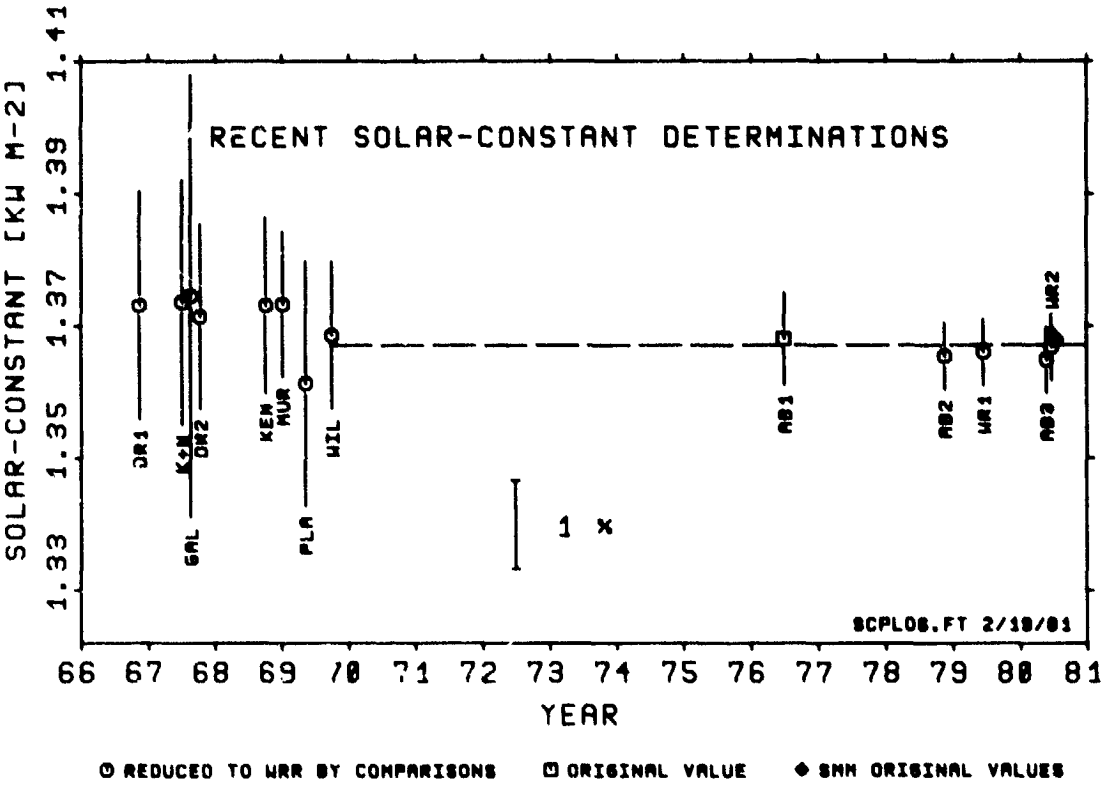


Figure 1: Recent solar constant determinations: DR1: Airplane, Drummond et al (1968); K+N: Balloon, Kondratyev and Nikolsky (1970), GAL: Airplane, Thekaekara et al (1969); DR2: X-15 Rocket aircraft, Drummond et al (1967); KEN: Airplane, Kendall (1973); MUR: Balloon, Murcray et al (1969); PLA: Mariner Satellite, Plamondon (1969); WIL: Balloon, Willson (1973); AB1: Rocket, Duncan et al (1977); AB2: Rocket, Willson (1980); WR1: Balloon, Brusa and Fröhlich (1980); AB3: Rocket, Willson (1980); WR2: Balloon, Brusa and Fröhlich (1981), SMM: Satellite, Willson (1981). (Detailed references can be found in ref. 7 and 8).

No significant change can be detected from the results of the period from 1969 to now: within ± 2 Wm^{-2} or ± 0.15 % the solar constant stayed constant. Although the values for the period before 1969 suggest a slight decrease, the uncertainties of these results are too large to render such a

conclusion significant. However, one could speculate and try to explain the decrease of the mean global temperature since the forties with a change of 0.1° per decade, by a decrease of the solar constant of 0.05% per decade, taking the data from 1966 to 1980 of figure 1.

VARIABILITY OF THE SOLAR SPECTRUM

Neckel and Labs (ref. 10) show that the accuracy of their revised spectrum is such that it can be used as a reference or baseline for the sixties, when their measurements were performed. Unfortunately, no other determinations with a similar accuracy and coverage have been made in the meantime. Only some values at a few distinct wavelengths have been determined with sunphotometers from Mauna Loa in 1978 (ref. 11). These sunphotometers have been calibrated either by the laser-radiometer method developed by Geist (ref. 13) or against irradiance standard lamps. Comparisons with the original and revised Neckel and Labs data are shown in figure 2. The agreement is within ±2%. This is excellent, if one takes into account the difficulties of such measurements and their calibration. It demonstrates also the improvement of the Labs and Neckel spectrum by using more accurate limb-darkening data. But due to the uncertainties involved, no conclusions can be drawn about a possible change from 1965 to 1978.

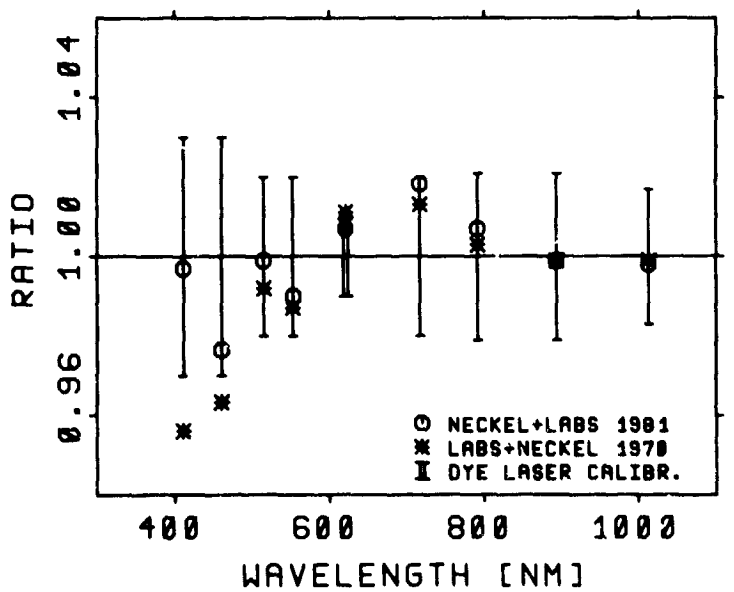


Figure 2: Comparison of Labs and Neckel (ref. 12), Neckel and Labs (ref.10) and Shaw and Fröhlich (Ref. 11) spectral irradiance data.

During the balloon experiment in June 1980, Ch. Wehrli of our institute has flown sunphotometers with centerwavelengths at 368, 500 and 778 nm and bandwidths of 5 nm. Preliminary results of the dye laser calibration of the 500 nm instrument yield a value of about 2% lower than the corresponding Neckel and Labs value. The value is just within the limits of the stated uncertainties and no conclusion about any longterm change can be drawn. On the other hand, the precision of the sunphotometers is so good that they can be used to monitor the solar output with a very high resolution during the few hours of the balloon flight. The result for the shortterm variability is shown in figure 3. Most important is the high correlation between the three wavelengths, which indicates the solar origin of the variability, because the three instruments are operated separately. To further exclude the possibility

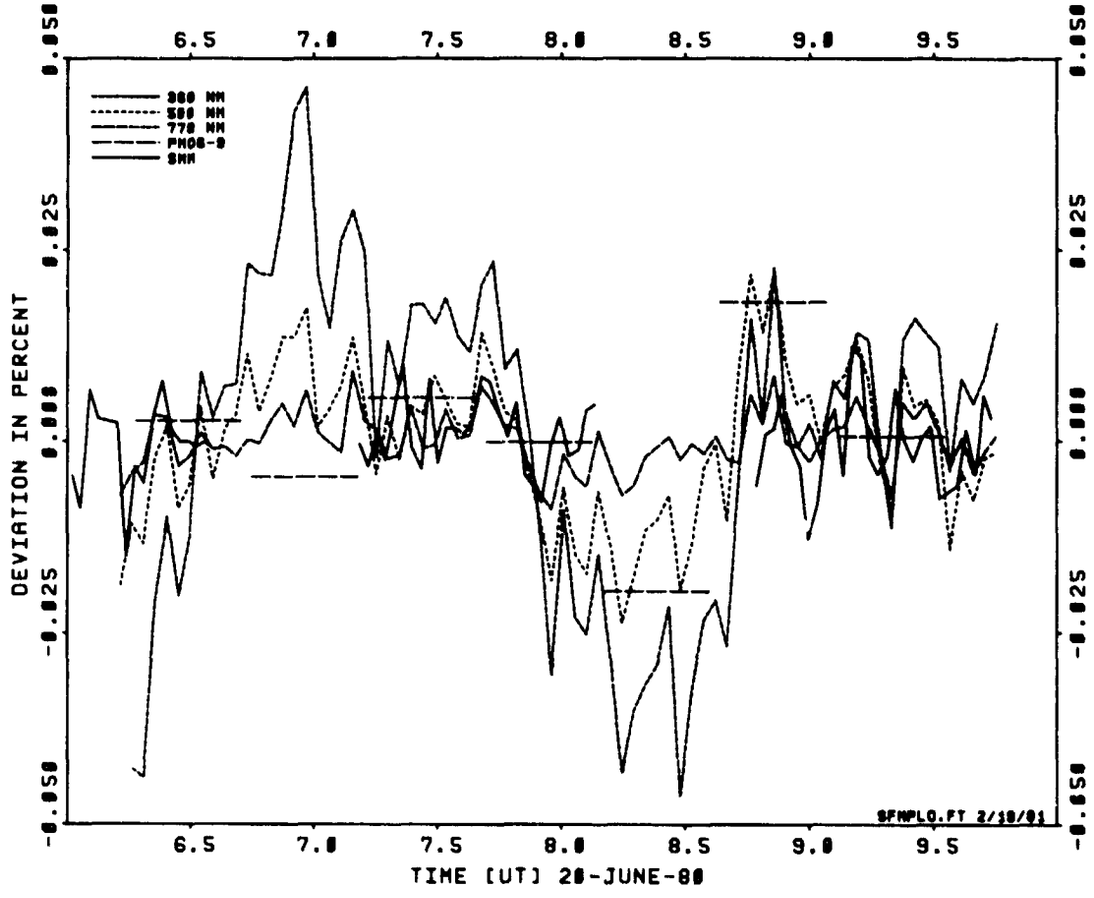


Figure 3: Shortterm variations of the solar output measured from a balloon at an altitude of 34 km with sunphotometers at 368, 500 and 778 nm and an absolute radiometer PMO6-9, compared to ACRIM radiometer data of SMM satellite.

of an instrumental, tracking or atmospheric influence, these results are compared with simultaneous measurements from the Solar Maximum Mission (SMM) total irradiance sensor ACRIM (*). Due to orbital restrictions, the SMM results cover only part of our records and unfortunately, only those with the least variation. But they do correlate very well, so that the solar origin of the variations is highly probable. Also the amplitudes compare well: as the solar spectrum is divided into two energetically equal parts at a wavelength of about 725 nm, the ACRIM amplitude is expected to lie somewhere between the amplitude of the 500 and the 778 nm record, which it does. Further analysis of these results are in course, but already these preliminary data reveal with a high degree of confidence for the first time the spectral dependence of the solar variability: the amplitude at 368 nm is more than double the amplitude at 500 nm, which in turn is about 1.5 times the amplitude at 778 nm or about 1.2 times the amplitude of the total irradiance variation measured from SMM.

The power spectrum analysis of the data show three major peaks for all three wavelengths at periods of 19, 6.5 and 5.2 minutes (figure 4). The last

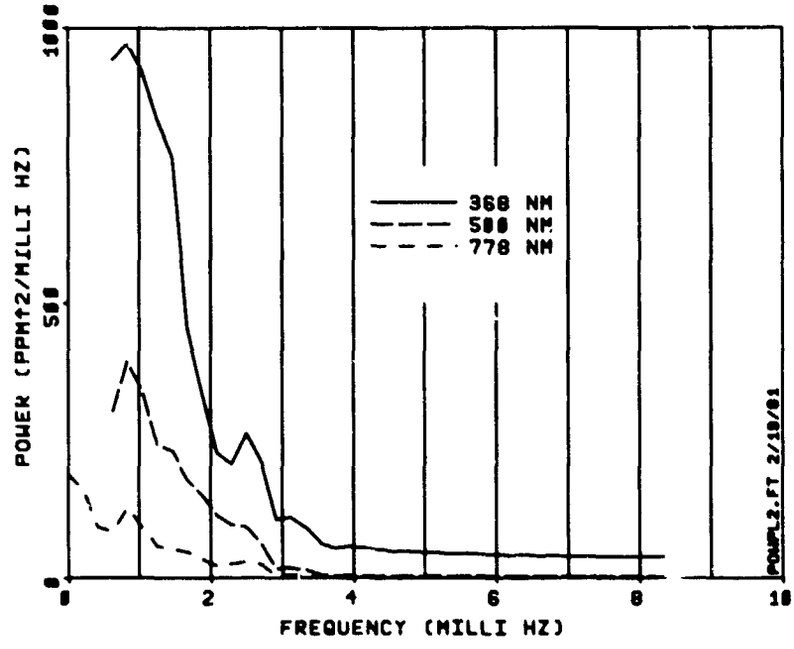


Figure 4: Powerspectrum of the shortterm variation of the solar output at 368, 500 and 778 nm.

(*) Willson, R.C.: Preliminary data of SMM ACRIM, private communication, 1980.

peak, although very weak, may indicate the intensity variation due to the 5-minutes solar oscillation. However, the sampling of two minutes is just marginal to detect this kind of variation. At least some of the variability could be due to the high activity of the sun during this period and thus bury more or less the amplitude of the 5-minutes oscillation.

CONCLUSIONS

The results of the analysis of the solar constant determinations reveal no indication for any significant change from 1966 to 1980. Unfortunately no data exist for the period from 1969 to the solar minimum. Hence, the question still remains, whether the total solar irradiance is changing with the solar activity cycle. The analysis of the available spectral data yields the same result but with less confidence, due to larger uncertainties.

A shortterm solar variability has been detected with amplitudes in the order of 100 - 500 ppm, which show a strong wavelength dependence, with an increase of a factor of three from 778 to 368 nm. Thus the measurement of these shortterm variations and their spectral distribution is capable to yield important information about solar variability and its spectral behaviour.

REFERENCES

1. Eddy, J.A.: Climate and the Changing Sun; Climatic Change, 1,No.2, 1977, pp. 173-190

2. Hoyt, D.V.: Variations in Sunspot Structure and Climate; Climatic Change, 2, No.1. 1979, pp. 79-92

3. Sellers, W.D.: A New Global Climatic Model; J. of Appl.Meteorology, 12, 1973, pp. 241-254

4. Timothy, J.G.: The Solar Spectrum Between 300 and 1200Å; The Solar Output and Its Variations, Colorado Assoc.Univ.Press, Boulder, 1977,pp.237

5. Heath, D.F. and Thekaekara, M.P.: The Solar Spectrum Between 1200 and 3000Å; The Solar Output and its Variations, Colorado Assoc.Univ.Press, Boulder, 1977, p. 193

6. Simon, P.C.: Solar Irradiance Between 120 and 400 nm and its Variatons; Solar Physics, 1981, in press

7. Fröhlich, C: Contemporary Measures of the Solar Constant; The Solar Output and its Variations, Colorado Assoc.Univ.Press, Boulder, 1977 p. 93

8. Fröhlich, C. and Brusa, R.W.: Solar Radiation and its Variation in Time; Solar Physics, 1981, in press

9. Willson, R.C.: Solar Total Irradiance Observations by the Active Cavity Radiometer; Solar Physics, 1981, in press

10. Neckel, H. and Labs, D.: Improved Data of Solar Spectral Irradiance; Solar Physics, 1981, in press

11. Shaw, G.E. and Fröhlich, C.: The Variability and Absolute Magnitude of Solar Spectral Irradiance; Solar-Terrestrial Influences on Weather and Climate, D. Reidel Publ.Company, Dordrecht, 1979, pp. 69-73

12. Labs, D. and Neckel, H.: Transformation of the Absolute Solar Radiation Data into the "Internat.Practical Temp.Scale of 1968"; Solar Physics, 15, 1970, pp. 79-87

13. Geist, J. et al: Spectral Radiometry: A New Approach Based on Electro-Optics; NBS Technical Note No. 954, 1977

A SUMMARY OF RESULTS FROM SOLAR MONITORING ROCKET FLIGHTS

Charles H. Duncan
Goddard Space Flight Center

ABSTRACT

Three rocket flights to measure the solar constant and provide calibration data for sensors aboard Nimbus 6, 7, and Solar Maximum Mission (SMM) spacecraft have been accomplished. The values obtained by the rocket instruments for the solar constant in SI units are: 1367 wm^{-2} on 29 June 1976; 1372 wm^{-2} on 16 November 1978; and 1374 wm^{-2} on 22 May 1980. The uncertainty of the rocket measurements is ±0.5%. The values obtained by the Hickey-Frieden (H-F) sensor on Nimbus 7 during the second and third flights was 1376 wm^{-2}. The value obtained by the Active Cavity Radiometer Model IV (ACR IV) on SMM during the flight was 1368 wm^{-2}.

INTRODUCTION

Three rocket flights to measure the solar constant and provide "ground truth" calibrations for spacecraft sensors have been accomplished to date. The first flight was initiated by NASA Headquarters in January 1976 because the values being obtained for the solar constant by the ERB flat-plate detector on Nimbus-6 were 1.5 percent higher than expected, i.e., 1392 wm^{-2}. Nimbus-6 first began taking data on July 2, 1975. This first flight identified a calibration error of +1.6 percent in the Nimbus-6 ERB channel 3 values. Subsequently, Hickey, et al., have identified the cause for +0.7 percent of this calibration error. The reasons for the remaining +0.9 percent error has not been identified to date. However, Hickey only claimed ±0.75 percent accuracy for this detector[1], so the values obtained from both spacecraft and rocket during the first flight are within the bounds of uncertainty of ±0.5 percent.[2] This paper provides some background on the history and results of these rocket flights.

FIRST ROCKET FLIGHT

Prior to authorizing the first rocket flight, NASA HQ convened an Ad Hoc Science Review Committee to consider the merits, probability of success and selection of experiment payload. The committee met initially on January 26, 1976. The personnel were:

Guenther Brueckner	–	Naval Research Laboratory, Washington, DC
Louis Drummeter	–	Naval Research Laboratory, Washington, DC
John Gille	–	National Center for Atmospheric Research, Boulder, CO
Verner Suomi	–	University of Wisconsin, Madison, WI

Robert Madden	--	National Bureau of Standards, Washington, DC
Jon Geist	--	National Bureau of Standards, Washington, DC

This committee made the following recommendations:

1. Ground intercomparisons at a high mountain site among rocket payload and neutral outside sensor [recommended PMO (Physikalisch-Meteorologisches Observatorium) Davos, Switzerland sensor developed by C. Fröhlich] be conducted prior to flight under ambient and vacuum conditions (10^{-4} Torr) and that agreement among all instruments be better than ±0.5 percent.

2. Tests to be conducted to verify behavior of payload upon exposure to the pressure temperature profile of the mission using thermal vacuum chambers and a solar simulator. Any effects upon performance to be noted and mission aborted if these effects cause more than 0.1 percent changes in response of individual instruments.

3. To minimize thermal problems and to eliminate all windows, payload to be launched in an evacuated configuration (10^{-4} Torr).

4. Instrument payload to consist of prototypes of ERB solar channels and as many self-calibrating radiometers as possible.

Subsequent to this meeting and prior to the flight, a final review of the results of intercomparisons, pressure-temperature profile testing, and related factors was held at NASA HQ on June 3, 1976. Upon presentation of the data, the Ad Hoc Committee gave its final approval for flight.

Pertinent test results presented at this meeting included:

1. Ratios of irradiance by the five payload instruments were shown to be constant at all irradiance levels at all pressures.

2. Calibration factors for pressure intermediate between 50 Torr and 10^{-4} Torr for ERB 3, ESP, and PACRAD were derived.

3. Simulation of launch pressure variations showed that all five instruments would read the solar constant within less than 45 seconds after first acquisition of the sun.

4. The five solar constant detectors (rocket payload plus two PMO detectors) agreed with each other within ±0.3% during the South Baldy intercomparisons at both ambient and vacuum (10^{-4} Torr).

5. An insect was trapped in canister upon placement of quartz window on payload at South Baldy and subsequently fell into the receiver of the PACRAD causing a 2.83 percent decrease in measurement data. Quartz was removed, bug was removed, PACRAD re-exposed to sun and obtained original results as compared to other instruments.

6. Payload instruments viewed LN_2 target (-185°C) at 10^{-6} Torr to determine zero offset.

Table I lists the instruments and investigators for each of the rocket flights. The first flight was launched on June 29, 1976 at 12:20 PM MDT from White Sands, New Mexico. During launch, according to readings obtained by the ERB-3 sensor (fast time constant, pressure sensitive) the canister lost vacuum during the initial launch phase and did not recover until the nose cone was blown. This fact was also verified by the PACRAD. The ACR IV, not having demonstrated pressure sensitivity during five track tests, did not note this fact.

The initial values reported by the investigators at about 3 minutes into flight did not change by more than 1 wm^{-2} except for the ERB-3 value which was initially reported as 1374 wm^{-2} and subsequently changed to 1389 wm^{-2}. The justification being: ERB on Nimbus-6 when pointed to

TABLE I. SOLAR MONITORING ROCKET FLIGHTS

- PAYLOAD FOR 1ST FLIGHT

INSTRUMENT	INVESTIGATOR/ INSTITUTION	SUPPORT CONTRACTOR
ACR IV 402A ACR IV 402B	R. C. Willson/JPL	TRW, Los Angeles, CA[1]
PACRAD	J. Kendall/JPL R. Harrison/JPL	None
ERB-ESP Channels 2 3 4 5 of Nimbus-6 ERB	J. Hickey/Eppley	Gulton Industries,[1] Albuquerque, NM

- CHANGES FOR 2ND AND 3RD FLIGHTS

  - ERB-ESP – Eliminate Channels 2, 4, and 5
    - Add Channel 3 With Anodized Baffles
    - Add H-F Sensor

10/27/80

[1]For Flights 1 and 2 only; no support contractors for third flight.

space produced a negative count equal to 15 wm^{-2}. If this negative offset were applied to the rocket data then 1374 +15 = 1389 wm^{-2} which made perfect agreement between rocket 1 and Nimbus-6 ERB. Table II lists the values obtained by the rocket instruments and Nimbus 6 ERB.

SECOND ROCKET FLIGHT

The program lay dormant from the first flight on June 29, 1976, until October 27, 1977, when an AN proposal was accepted for four additional rocket flights to be conducted under the general philosophy developed by the Ad Hoc Science Review Committee.

The payload was refurbished; Hickey substituted an ERB-3 channel with anodized baffles and an improved version of the ESP, termed an H-F, for the filter solar channels flown on the first flight. These channels did not obtain any data which gave meaningful insight into the behavior of the same channels on Nimbus-6 ERB. Each of these channels was covered with a filter and some values were higher and some lower on the rocket as compared to the spacecraft with differences of 27 wm^{-2}, 20 wm^{-2}, and 1 wm^{-2} noted.

The addition of the ERB-3 with anodized channels was added to try to identify the source of the calibration error of Nimbus-6 ERB discovered by the first rocket flight. The H-F sensor was added because it had become part of the payload of Nimbus-7 as channel 10 C.

The intercomparison of the Nimbus-7 channel 10 C detector with the rocket payload prior to the lau ch of Nimbus-7 could not be accomplished. Subsequently, the second rocket flight was launched on the same day and obtained values at the same time as the Nimbus-7 ERB was first turned on. Also, during the second flight, the payload was pointed off sun for 30 seconds to try to verify that the space offset of ERB-3 was really 15 wm^{-2}. The payload also lost vacuum about 15 minutes prior to launch.

The primary result from the second flight was an apparent increase in the solar constant since the first flight. The result of off sun pointing was inconclusive since the ERB detector first began exhibiting negative counts, then reversed this trend, then became negative again for a few seconds, then acquired the sun. This effect was most probably due to the fact that the nose cone was drifting into and out of the field of view of the instrument. The negative offset maximum observed was equivalent to approximately 12 wm^{-2}. This verifies that the space offset observed for Nimbus 6 was most probably the same for the rocket although sufficient observation time was not available to reach 15 wm^{-2}. It should be noted that: according to Hickey[1] "The space look offset amounts to almost 0.9 percent of the solar constant value for ERB channel 3."

The most disappointing aspects of this flight were the saturation of the H-F sensor in flight due to the wrong value of heater power being set, the saturation of the ACR IV-B, and the initial low values, 2-3 percent lower than expected, obtained by the ESP. The ESP lower values were later

[1] Private communication.

TABLE II. SOLAR MONITORING ROCKET FLIGHTS

Results of 1st Flight – June 29, 1976 @ 12:20 PM MDT – White Sands, NM

INSTRUMENT	VALUE
PACRAD	1364 wm^{-2}
ACR A	1368 wm^{-2}
ACR B	1368 wm^{-2}
ESP	1369 wm^{-2}
Mean	1367 wm^{-2}
ERB Rocket Channel 3	1389 wm^{-2}
ERB Nimbus Channel 3	1389 wm^{-2}
Difference ERB Channel 3 and Rocket Payload Mean	+22 wm^{-2} or +1.6%

10/28/80

found to be due to aperture contamination and a change in the parameters relating back heater power to irradiance. A recharacterization of the ERB sensors accomplished during this activity discovered a +0.7% calibration error for the Nimbus-6 ERB sensors as flown on Nimbus-6 and the first rocket flight. The anodized baffles on ERB-3 gave the same results as the painted baffles so this channel yielded no new information to resolve the calibration anomaly noted on the first flight. Table III lists the values for this flight along with the Nimbus 6 & 7 ERB values. The values shown for the ESP include the convection of 1.96% for aperture contamination and 0.3% for heater power parameter changes. Both values are shown.

Because of the importance of the possible increase in the solar constant, a post-flight intercomparison was held at Table Mountain, California, on 10-13 December 1978. Hickey was unable to participate in this intercomparison. However, Fröhlich of WMO and Crommelynck of Belgium both participated. Crommelynck has been selected to fly a solar constant experiment aboard the shuttle and brought his prototype for this flight to Table Mountain. Fröhlich brought with him the same instruments which had initially been used at the South Baldy intercomparisons. The results of this intercomparison were that the PMO and ACR IV-A were still reading within 0.06 percent of the intercomparison values at South Baldy while the PACRAD was reading 0.5 percent higher. Kendall and Harrison reported subsequently[2] that an incorrect characterization had been used in the South Baldy intercomparisons, but the correct characterization had been used for both the first and second rocket flights and for the post-flight intercomparisons, after the second rocket flight. They maintained that their instrument had truly shown an increase in the solar constant between the two flights.

During this time period, Willson intercompared his three sensors flown on the SMM with the rocket ACR IV instruments. In comparison to the rocket ACR IV-A, the SMM sensors A, B, and C read 0.2 percent, 0.04 percent, and 0.3 percent higher respectively. The Crommelynck sensor read 0.7 percent lower, the PMO2 read 0.5 percent lower and the PMO6 read 0.15% lower than the ACR 402A.

The value derived from the second rocket flight for the solar constant of 1372 wm^{-2} was 4 wm^{-2} lower than the value obtained simultaneously by Nimbus-7 channel 10 C (H-F). However, the rocket instruments were not intercompared with the Nimbus-7 channel 10 C before launch of Nimbus 7 so the bias between the rocket instruments and the H-F could not be identified since the rocket H-F saturated. These two values were within the uncertainty of the measurements however.

THIRD ROCKET FLIGHT

As a result of the values obtained from the second flight, the investigators (Willson, Hickey, Kendall, Harrison) took extreme care in the preparation and execution of the third flight.

Extensive intercomparisons before the flight were made on 15-19 April 1980. During this intercomparison, the relative performance of the rocket sensors was established again. The

[2]Private communication.

TABLE III. SOLAR MONITORING ROCKET FLIGHTS

Results of 2nd Flight – November 16, 1978 @ 11:15 MST – White Sands, NM

INSTRUMENT	VALUE
PACRAD	1371 wm^{-2}
ACR A	1373 wm^{-2}
ACR B	Saturated
ESP	1373 wm^{-2}* – 1378 wm^{-2}
H-F	Saturated
Mean (No-ESP)	137[illegible] wm^{-2}
Mean (With-ESP)	1372–1374 wm^{-2}
ERB Rocket Channel 3	1383 wm^{-2}
ERB Nimbus-6 Channel 3	1387 wm^{-2}
ERB Nimbus-7 Channel 3	1383 wm^{-2}
Mean	1384 wm^{-2}
Nimbus-7 H-F Channel 10C	1376 wm^{-2}
Difference ERB Channels 3 and Rocket Payload Mean	+12 wm^{-2} or +0.9%
Difference H-F Channel 10C and Rocket Payload Mean	+4 wm^{-2} or +0.3%

*Value derived after disassembly of sensor.

PACRAD and ACR IV were consistent with their performance during the initial intercomparisons at South Baldy peak. The ESP read about 0.3% higher than during the intercomparisons before the first flight. The H-F which had never been intercompared previously read about the same as the PACRAD; about 0.3 percent lower than the ACR's, and about 0.8 percent lower than the ESP.

The decision was made prior to the third rocket flight not to try to evacuate the canister since vacuum had not been maintained after launch for either of the two previous flights and the effects of dynamic heating were negligible on causing temperature excursions on the instruments.

Consequently, the payload was purged with dry air for a week before launch whenever possible and continuously until launch after the horizontal test was complete, a period of about 48 hours.

The third flight had been scheduled to coincide with the first turn-on of SMM ACRIM. Problems with the Aerobee 170 rockets forced a delay of the flight to May 22, 1980 after a decision had been made to fly the payload aboard an Astrobee rocket instead of continuing to wait for an Aerobee 170 to be readied.

The results from the third flight agree very closely with those from the second flight for both rocket instruments and Nimbus-7 H-F measurements. The results are summarized in Table IV. Nimbus-7 channel 10 C measured 1376 wm^{-2} for the dates of each of these flights and the rocket averages for the second and third flights were 1372 and 1374 wm^{-2} respectively, all well within the estimated uncertainty of the instruments.

However, the SMM ACRIM was also in space on 22 May 1980 obtaining a value of 1368 wm^{-2} for the solar constant. These ACRIM instruments (now reading lower in space) had read about 0.2 percent higher than the rocket ACR IV's during intercomparisons at Table Mountain in December 1979. This result indicated that environment, possibly pressure, might influence the measurements of the ACR detectors. Also, the H-F during intercomparisons in April 1980 at Table Mountain read 0.3 percent lower than the rocket ACR but during the third rocket flight, it read 0.3 percent higher for a total difference of 0.6 percent. The rocket H-F, however still read about 0.2 percent higher than the Nimbus-7 H-F for a total difference between ground intercomparisons and spacecraft values of about 0.4 percent which is very close to the difference (0.6%) observed by Willson between ground, rocket, and space performance.

SUMMARY

Table V lists pertinent data for each of the three flights along with sunspot numbers on the dates of the flights. Table VI summarizes the results of the measurements for the three flights. Table VII summarizes the results from Nimbus 6 & 7 ERB and SMM ACRIM for the dates of the flights. Table VIII presents an average value for the solar constant for the dates of each flight using both rocket and spacecraft data. The rocket average value for the solar constant is given equal value to the spacecraft values in a simple arithmetical average.

TABLE IV. SOLAR MONITORING ROCKET FLIGHTS

Results of 3rd Flight – May 22, 1980 @ 9:00 MDT – White Sands, NM

INSTRUMENT	VALUE
PACRAD	1373 wm^{-2}
ACR A	1373 wm^{-2}
ACR B	1374 wm^{-2}
H-F	1378 wm^{-2}
ESP	1385 wm^{-2}
Mean	1377 wm^{-2}
ERB Rocket Channel 3	1377 wm^{-2}
ERB Nimbus-6 Channel 3	1377 wm^{-2}
ERB Nimbus-7 Channel 3	1367 wm^{-2}
Mean	1374 wm^{-2}
Nimbus 7 H-F Channel 10C	1376 wm^{-2}
SMM ACRIM A, B, C Average	1368 wm^{-2}
Difference ERB Channels 3 and Rocket Payload Mean	3 wm^{-2} or 0.2%
Difference H-F Channel 10C and Rocket Payload Mean	1 wm^{-2} or 0.07%
Difference SMM ACRIM and Rocket Payload Mean	9 wm^{-2} or 0.6%

TABLE V. DATA RELATIVE TO ROCKET FLIGHTS

Date	Time of Launch	Peak Altitude	Length of Obs. Int'v.	Start of Meas'mt. Int'v.	Zurich SS Number	Ottawa 2808 MHz Flux	Bartels 27 Day Cycle No.
29 June 76	1220 MDT	255 km @ +247 sec.	311 sec.	+ 93 sec.	11	71.3	7
16 Nov. 78	1115 MST	183 km @ +222 sec.	212 sec.	+133 sec.	77	128.8	13
22 May 80	0900 MDT	249 km @ +260 sec.	339 sec.	+ 94 sec.	244	276.6	26

TABLE VI. VALUES OF THE SOLAR CONSTANT FROM ROCKET FLIGHTS

(All Values in wm^{-2})

Date	ACRA	ACRB	PACRAD	ESP	II-F	ERB-3 Painted	ERB-3 Anodized	Average Value
29 June 76	1367	1368	1364	1369	[3]	1380	[3]	1367
16 Nov. 68	1373	[1]	1371	1373 – 1378 [2]	[1]	1383	1381	1372
22 May 80	1373	1374	1373	1385	1378	1377	1374	1374 [4]

Notes:
[1] Saturated during flight.
[2] See text for explanation.
[3] Not flown on this flight.
[4] ESP not included in average value.

TABLE VII. SPACECRAFT SOLAR CONSTANT VALUES AND SUNSPOT NUMBERS FOR DATES OF ROCKET FLIGHTS

Date	Zurich Sunspot Number	Nimbus-6 (ERB-3) Channel 3	Nimbus-7 (ERB-3) Channel 3	Nimbus-7 Channel 10-C (H-F)	SMM (ACR) ACRIM	Mean of Rocket Values
29 June 76	11	1389	N.A.	N.A.	N.A.	1367
16 Nov. 78	77	1387	1383	1376	N.A.	1372
22 May 80	244 (provisional)	1377	1367	1376	1368	1374

TABLE VIII. AVERAGE VALUE OF SOLAR CONSTANT AT TIMES OF ROCKET FLIGHTS (CAVITY SENSORS IN SPACE)

Date	Sensors	Value (wm^{-2})	Average Value (wm^{-2})
29 June 1976	Rocket Payload	(1367)	1367
16 Nov. 1978	Rocket Payload	(1372)	1374
	Nimbus 7 H-F	(1376)	
22 May 1980	Rocket Payload	(1374)	1373
	Nimbus 7 H-F	(1376)	
	SMM ACRIM	(1368)	

A fourth rocket flight is planned for May 1981. The experiment canister is being rebuilt so that vacuum can be maintained prior to and after launch. Fröhlich will also provide two PMO-6 sensors for the experiment payload under an international agreement which has been negotiated recently. Plans for this flight also include 30 seconds off sun pointing after 5 minutes of data have been taken to establish the space offset values for the sensors.

REFERENCES

1. J. R. Hickey and A. R. Karoli, Appl. Opts., 13, 523 (1974).

2. C. H. Duncan et al., Appl. Opts., 16, 2690 (1977).

SOLAR VARIABLITY

INDICATIONS FROM NIMBUS 7 SATELLITE DATA

J.R. Hickey, B.M. Alton, F.J. Griffin
The Eppley Laboratory, Inc.

H. Jacobowitz, P. Pellegrino
NOAA/National Earth Satellite Service

E.A. Smith and T.H. Vonder Haar
Colorado State University

R.H. Maschhoff
Gulton Industries

ABSTRACT

The cavity pyrheliometer sensor of the Nimbus 7 Earth Radiation Experiment (ERB) has indicated low-level variability of the total solar irradiance. The variability appears to be inversely correlated with common solar activity indicators in an "event" sense. The limitations of the measuring system and available data sets are described.

FOREWORD

The content of this paper is modified from the presentation delivered at the workshop. Some new data and subsequent discussion has been added. Much of the background information has been deleted but is referenced. The figures have been updated to the latest possible available time depending on the data set. The reprocessing effort for the Nimbus 6 data has begun at the time of this writing (early February 1981); however, insufficient information is available to update the Nimbus 6-7 overlap agreement. Only the results from the cavity sensor of Nimbus 7 are presented here.

INTRODUCTION

Solar parameter measurements have been performed since November 1978 by a self-calibrating cavity pyrheliometer on the Nimbus 7 satellite. The results presented here must be

considered preliminary because of the nature of the data sets employed. These data sets and their limitations are described in the opening paragraphs of the paper. Correlation with other solar activity indicators are presented as well as a comparison with results from the Solar Maximum Mission.

CALIBRATED DATA (SEFDT)

The orbital values are available for the months of November and December 1978 and January, February, March, June and October of 1979. This data is designated SEFDT data based on the source tape. These data have been processed to the highest degree possible. That is, the correct earth-sun distance, temperature correction, and space-offset have been applied to the data set. The in-flight calibration factors have been applied as obtained from analysis of the heater calibration sequences which are available from an independent data source. The SEFDT set comprises 135 daily mean values for the period. The missing days within the months listed above are not available nor will they be. This is because of the operational schedule of Nimbus 7. The missing months are due to the fact that the processing has not been performed for those periods.

Figure 1 is a plot of the available daily means from this calibrated data set versus time. The solar indicators of sunspots and 2800 MHz flux are plotted on the same time axis for comparison. The mean value is 1374.25 Wm^{-2}. The standard deviation is 0.623 Wm^{-2} or 0.0453% of the mean. The range of irradiance values is from 1372.79 to 1375.45 or 0.19% of the mean. The range is approximately 4.3 times the standard deviation. The minimum value is about 0.11% below the mean while the maximum is about 0.09% above the mean. The slope of the regression shows a small downward trend of -0.9 Wm^{-2}/1000 days. This amounts to -0.065%/1000 days or -0.024%/year.

It is probable that the absolute value of the mean irradiance is too high. Recent investigation of available off-axis flight data indicates that the stray light correction in the calibration equation is underestimated by about 0.2%. While further investigation is required to confirm this effect, it is likely that this correction should be applied. The corrected mean value would then be 1371.5 Wm^{-2}.

ENGINEERING DATA

This data set covers the period from November 16, 1978 through January 5, 1981. The set comprises daily mean values

for the period through February 4, 1978 but only one orbit per day for the remainder of the data. The basic sensor output counts are obtained as a peak signal output from the engineering analysis program at the Nimbus ground station. The space-offset value, temperature correction, earth-sun distance adjustment and calibration are applied to the data. A description of the limitations of this data set were presented (1) by the ERB team in describing the early results. While there is additional information available now from the analysis of the calibrated (SEFDT) data set, the engineering data has been continously processed in a consistent manner. The absolute value of irradiance is generally higher for the engineering set because of the method of application of the space-offset term and the temperature correction. The former is because it is either not present in the data or because it is available for the orbit minimum only as a single point value. The latter is because the pre-flight rather than the in-flight derived coefficient is employed.

Since both the space-offset and on-sun signals are single points and since the uncertainty of a single point is $\pm$ 0.5 count of digital output, the uncertainty is $\pm$ 1 digital count for this reason alone. For mean earth-sun distance, this is equivalent to about 0.056 percent or 0.77 Wm^{-2}. There are 463 daily values in the engineering data set for the period of 782 days. The reasons for missing data here are the same as for calibrated data with the additional reason that the engineering analysis program output is not available for every day of ERB operation.

Figure 2 is a plot of the engineering data set for the period. In this plot all of the points have been connected despite the fact that it is not a continuous data set. This figure is included to give the reader a feeling for variability over the entire period. The mean value is designated by the horizontal line at 1375.55 Wm^{-2}. The regression line shows a very small downward trend over the period which can be expressed as -1 Wm^{-2} per 1000 days or 0.073%/1000 days or approximately 0.026%/year. The standard deviation is about 0.071% of the mean value. This is slightly larger than the uncertainty of a single value (0.056%) discussed previously. The range of the data is 0.443% of the mean with the minimum value at -0.27% and the maximum value at +0.17%. Thus, the low value is almost 4 sigma below the mean while the high value is about 2.5 sigma above it. The region of low irradiance values in August of 1979 is the most prominent feature on the plot.

Figures 3a and 3b show the same data set split into two 400 day periods. The points are not connected so that the

discontinuities are recognizable. The solar activity indicators of sunspot number and 2800 MHz flux are also plotted on the same time scale for comparison.

COMPARISON OF DATA SETS

There are only 108 points common to the SEFDT and Engineering data sets. It must be remembered that the early engineering data are daily means while only one point per day is available after February 4, 1979. There are 44 points which are daily means for both sets in this discussion. The mean ratio of engineering to SEFDT data is 1.00116 indicating that engineering data is higher on the average by 0.116%. This is equivalent to two counts of input signal as discussed earlier. The nature of the agreement between the sets is shown in Figure 4. The correlation coefficient is only 0.528. The standard deviation is smaller for the SEFDT set (0.046%) than for the engineering data (0.061%) as would be expected. Lines indicating the means of both sets and the regression line are shown on the plot. It should be noted that a number of values fall in the second and fourth quadrants indicating opposite behavior of the data sets about their respective means. Most of these points are close to the quadrant separator lines within the expected uncertainties.

It is noted that the periods which show very low values in the engineering data are generally misssing from the available SEFDT data. Notably, the large August 1979 event and lesser events in September and November of 1979 are not included. Also, two events which were sensed by the SMM ACRIM (discussed later) are not included in the set. The plot is essentially bounded by $\pm$ 0.1% deviation on the calibrated data while the engineering data extends from -0.20% to + 0.12%. An analysis on an orbit-by-orbit basis may improve this correlation. The additional information necessary to identify corresponding orbits is not available to us at this time. A preliminary analysis of the SEFDT orbital values has indicated variability at the 0.02% to 0.07% range on a daily basis.

COMMENTS ON SOLAR VARIABILITY

It had been noted early in the project that dips in the engineering data corresponded to peaks in the 2800 MHz and sunspot data in an event sense. However, until higher quality data became available and until the large unambiguous dip occurred in August of 1979, it was not possible to rule out instrumental effects. The identification of a change was

reported (2) at the AGU meeting in May 1980 after a confirmation of the engineering data indications by a calibration orbit analysis. An interim set of "preliminary scientific" data obtained from "Master Archive Tapes" (MATS) was, also, employed in confirming the variability (2,3). This set of processed data was found to be improperly corrected for earth-sun distance and had other minor flaws none of which compromised its use in assessing short-term variability. Both the preliminary MAT results and the calibration orbit analysis (3,4) confirmed the short-term variability in the engineering data set although not the magnitude. The higher quality data of the preliminary SEFDT data set became available (4) which led to the previous discussion herein. In August 1980, we were supplied with the results of the ACRIM radiometer aboard the Solar Maximum Mission (SMM) satellite by Dr. R. Willson of JPL. These results have now been published (5) and were discussed at the workshop. A comparison of the SMM results to the FRB engineering data is shown in Figure 5 for 153 days starting in February 1980.

The circles represent the ERB data overlayed on Willson's original plot. The higher resolution of the SMM instrument is obvious from the plot. Both plots are expressed in percent deviation from the mean for the period. The two prominent dips in the SMM data correspond to the two most prominent dips in the ERB data. The larger of the two is about -0.16% for SMM and -0.21% for FRB. The other at day 147 is about -0.09% fro SMM and -0.18% for FRB. Thus, both the ERB deviations are greater than the SMM indications. The remainder of the plot shows correspondence within the engineering data uncertainty with a few notable high values on day 142 and in the period 165 to 180. We must wait for the calibrated FRB data before making conclusions relative to the remainder of the period or for evaluating differences in magnitude or corresponding events. It is noted that the large August 1979 dip was -0.27% below the mean.

From all of the correlative data it appears that the variability of the solar irradiance evident in the engineering data set is confirmed in principle but not necessarily in magnitude. All correlative data indicate lower magnitudes of the deviations. Unfortunately, no correlative data other than calibration orbits cover the periods of the greatest indicated variations.

CORRELATION WITH OTHER INDICATORS

A number of simple correlation analyses have been performed for the ERB data sets versus sunspots and 2800 MHz. These are

given below.

data set	correlation coefficient vs	
	sunspot Rz	2800 MHz
SEFDT	- .311	- .321
Engineering	- .284	- .342
Cal. orbit	- .441	- .550

The calibration orbit data represents only 43 points spaced over the period November 1978 through December 8, 1980. While these coefficients are not impressive, they do indicate an inverse correlation. The values of the coefficients are declining as each data set increases in number. In all cases, the 2800 MHz correlation has been higher than the sunspot coefficient. A preliminary event analysis was reported for earlier versions of the data sets (6). There was an indication that 16 of 25 identifiable events were unambiguously anti-correlated. An updated analysis was performed for a 654 day set of engineering data. The nine-day running mean of engineering data, 2800 MHz flux and sunspot number are shown as a function of time in Figure 6. It is easier to visualize the correspondence of events with this smoothed data. All depressions in the ERB data below 1374.5 Wm^{-2} coincide with peaks in the 2800 MHz flux and sunspot plots. This accounts for the 7 major dips in the smoothed data. Of the 13 peaks in the channel 10c which extend above 1376 Wm^{-2} 3 coincide with peaks in the 2800 MHz flux. While correlations can be found in the range of 1374.5 to 1376 Wm^{-2}, they must be interpreted with caution based on the previous discussion of the correlation between the engineering data and the calibrated data. If we consider every identifiable peak in the 9-day 2800 MHz data, there are 26 including many small bumps. Of these, 18 could be considered to coincide with dips in the ERB data. The two dips noted in the SMM data are identified by "x" on the vertical dashed lines.

We will not present further details of the correlation analysis here. We have not attempted to discuss the solar physics implications of these results. We have noted (4) that the large August 1979 dip coincided with the passage of a Coronal hole across the solar disc.

SUMMARY AND CONCLUSIONS

The current status of the ERB/NIMBUS 7 solar constant measurements have indicated that solar variability at the ± 0.1 to $+0.2$ level. The most probable value of the solar constant derived from the highest quality data and adjusted for underestimated reflection in the sensor is 1371.5 Wm^{-2}. The major depressions in the solar flux are correlated in an event sense with peaks in the sunspot numbers and 2800 MHz flux. Further detailed analysis awaits availability of a complete high quality data set.

REFERENCES

1. Hickey, J.R.; Stowe, L.L.; Jacobowitz, H.; Pellegrino, P.; Maschhoff, R.H.; House, F.B.; and Vonder Haar, T.H.: Initial Solar Irradiance Determinations from Nimbus 7 Cavity Radiometer Measurements. SCIENCE, 208, pp 281-283, 18 April 1980.

2. Hickey, J.R.; Griffin, F.G.; Jacobowitz, H.; Stowe, L.L.; Pellegrino, P.; and Maschhoff, R.H.: Comments on Solar Constant Measurements from Nimbus 6 and 7. Proceedings of the Spring Meeting of the American Geophysical Union (AGU): Toronto, May 1980.

3. Hickey, J.R.; Griffin, F.G.; Alton, B.M.; Jacobowitz, H.; Stowe, L.L.: Pellegrino, P.; and Maschoff, R.H.: Solar Parameter Measurements from Nimbus Satellites and other Solar Constant Experiments. Volume of extended abstracts of the IRS 1980 International Radiation Symposium, August 11-16; pp 343-345.

4. Hickey, J.R.; Griffin, F.J.; Alton, B.M.; Jacobowitz, H.; Stowe, L.L.; Pellegrino, P.; Maschhoff, R.H.; House, F.B.; Smith, E.A.; and Vonder Haar, T.H.: Review of Solar Irradiance Measurements from Nimbus 6 and 7. Proceedings (extended abstracts) of the Symposium at International Pyrheliometric Comparisons - IPC V, Davos, Switzerland, October 1980.

5. Willson, R.C.; Gulkis, S.; Janssen, M.; Hudson, H.S.; Chapman, G.A.: Observations of Solar Irradiance variability, SCIENCE, 211, pp 700-703, 13 February 1981.

6. Smith, F.A.; and Hickey, J.R.: Solar Variability Analysis Using Nimbus 7 ERB Results. (same as 4).

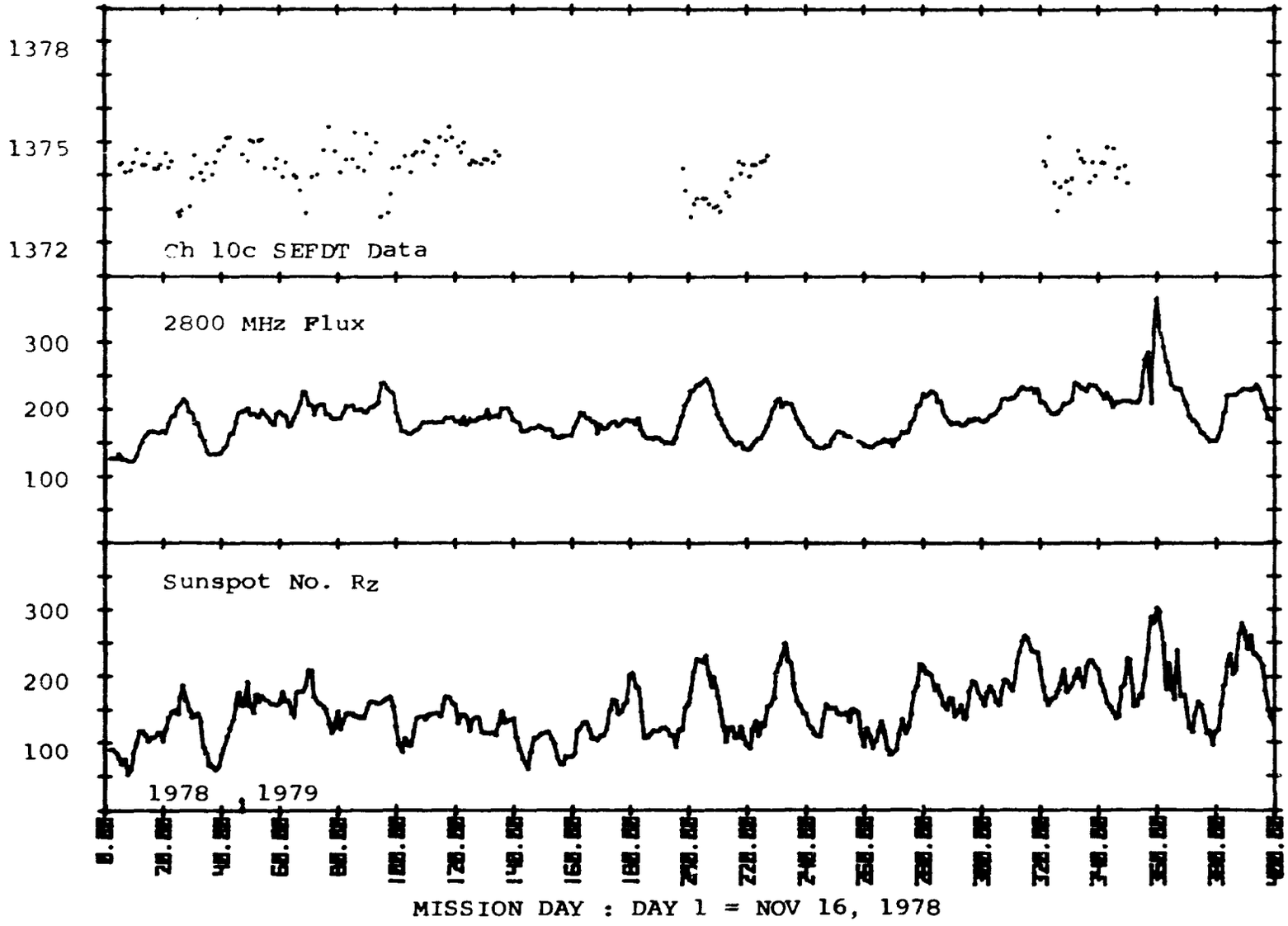


Figure 1. Calibrated ERB cavity data and solar indicators versus time

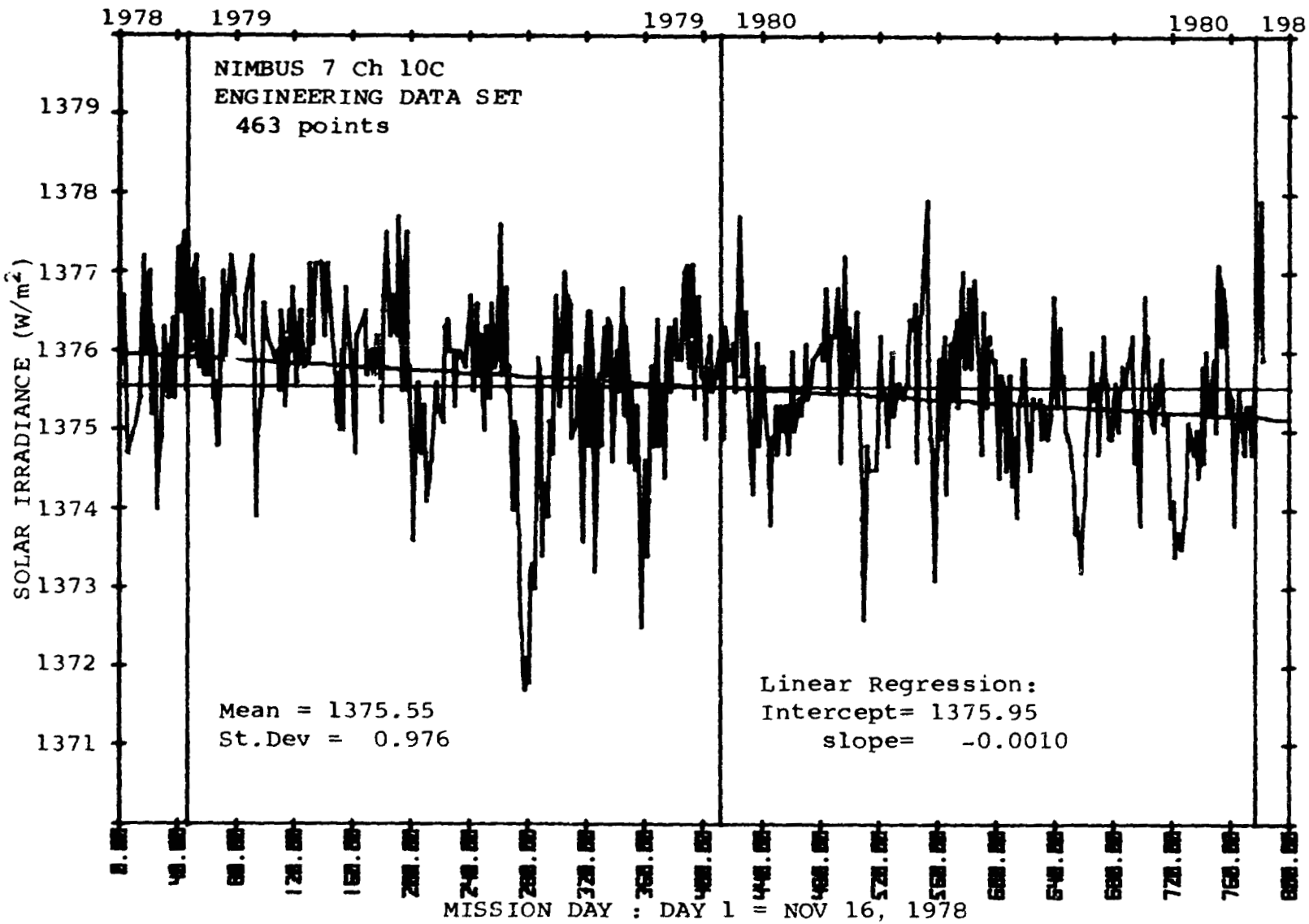


Figure 2. ERB channel 10c Engineering data versus Time in Nimbus-7 mission days

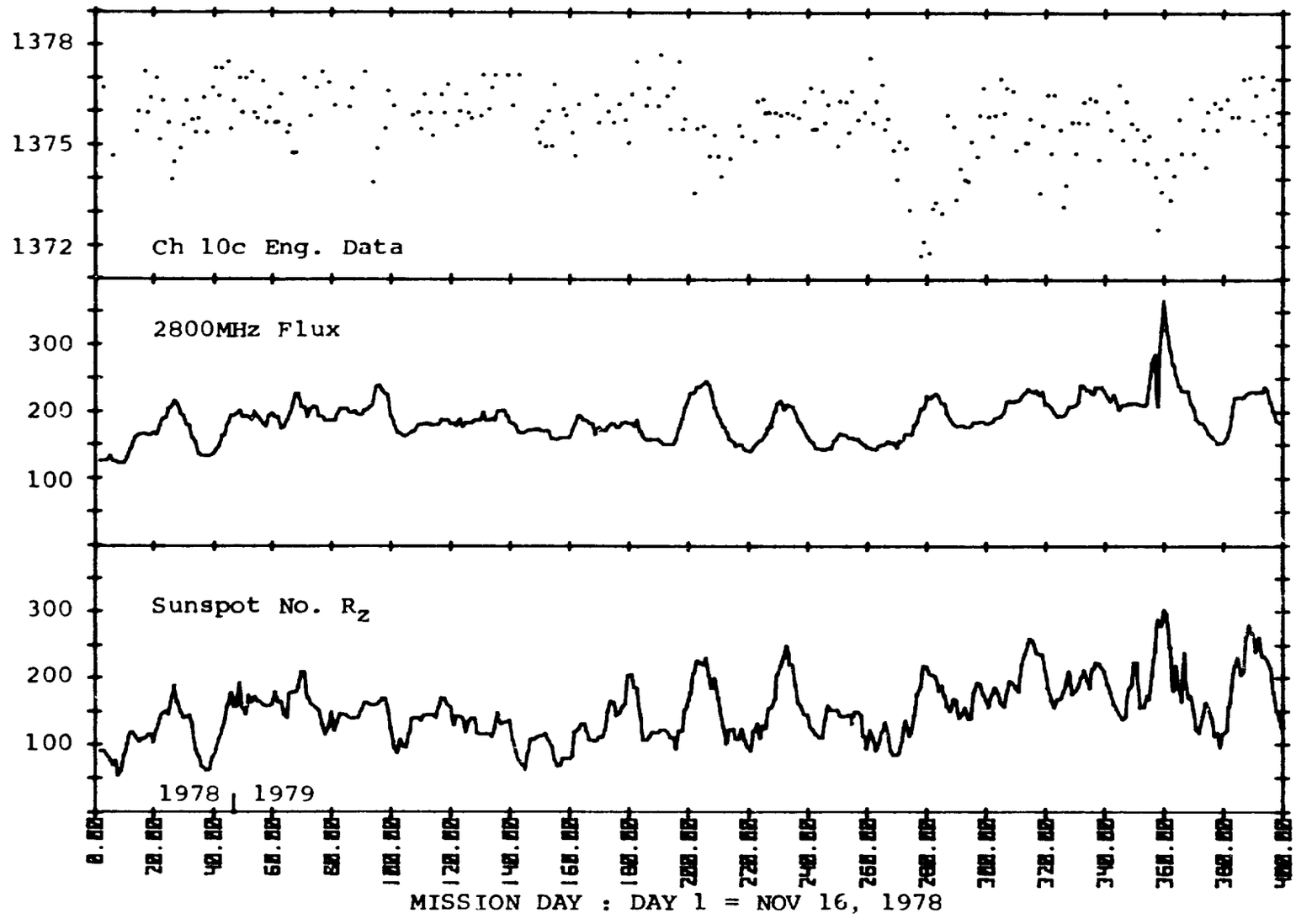


Figure 3a. ERB channel 10c engineering data and solar indicators versus time
First 400 days

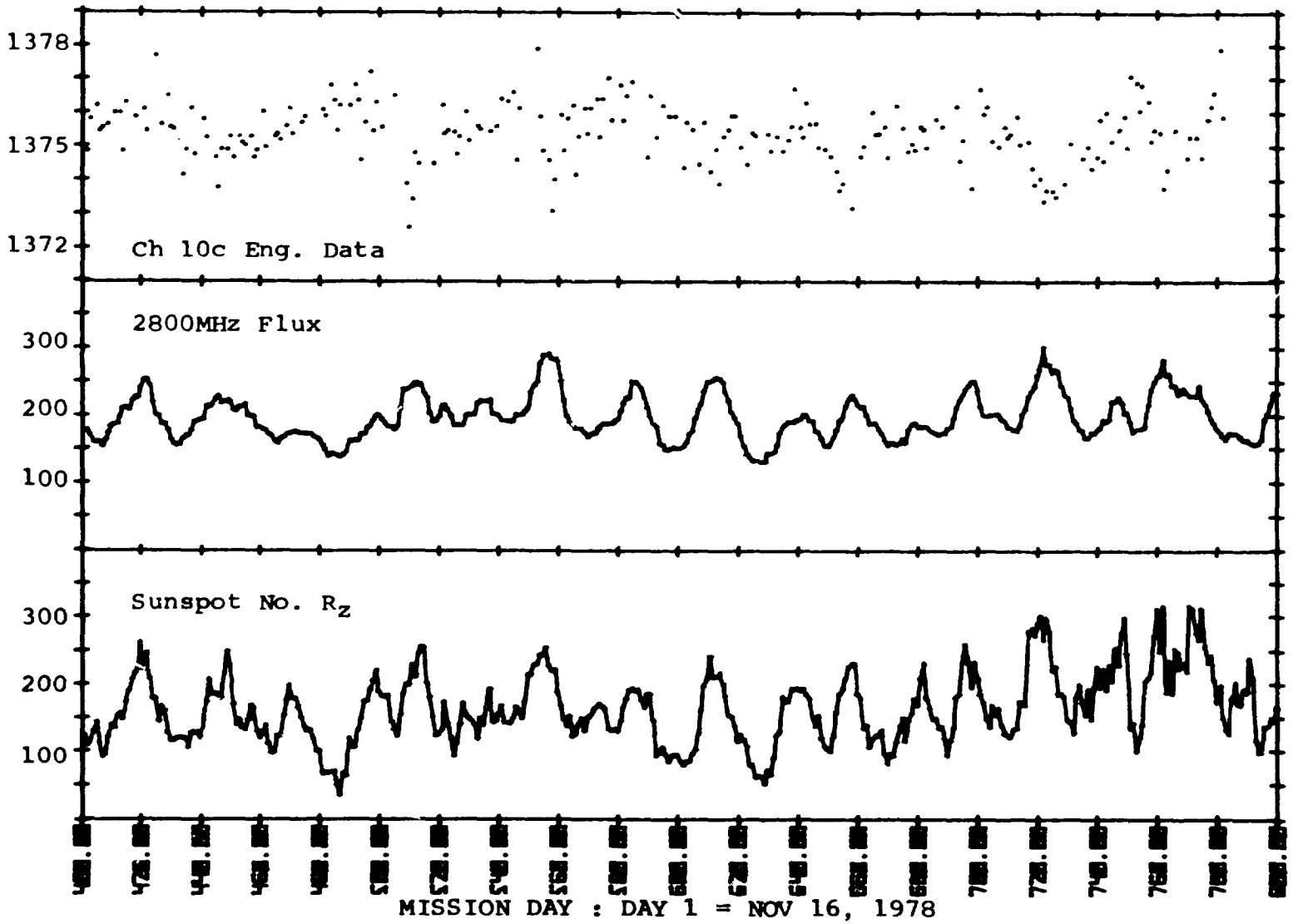


Figure 3b. Same as 3a. for days 401 through 782

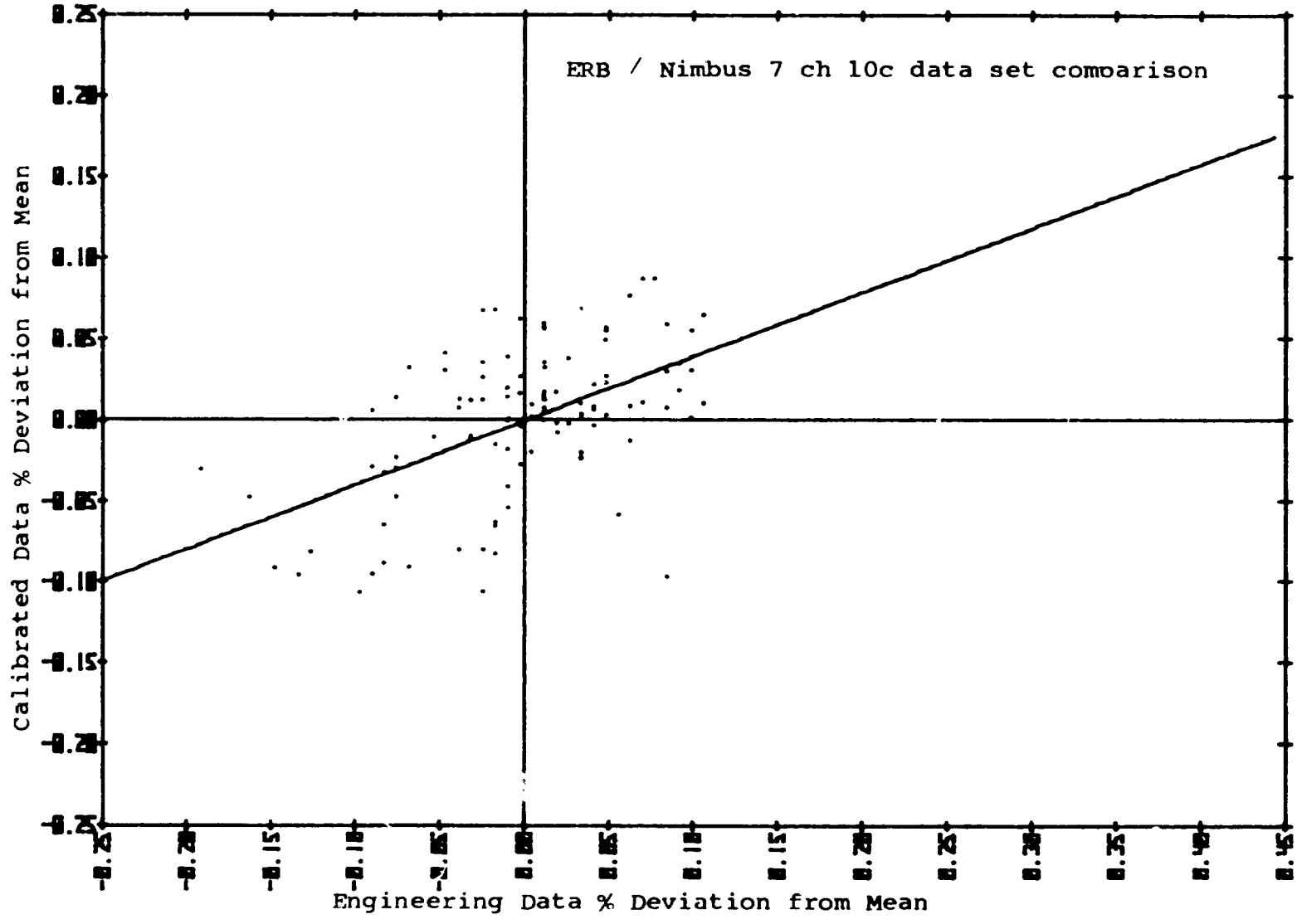


Figure 4. Comparison of deviations between calibrated and engineering data

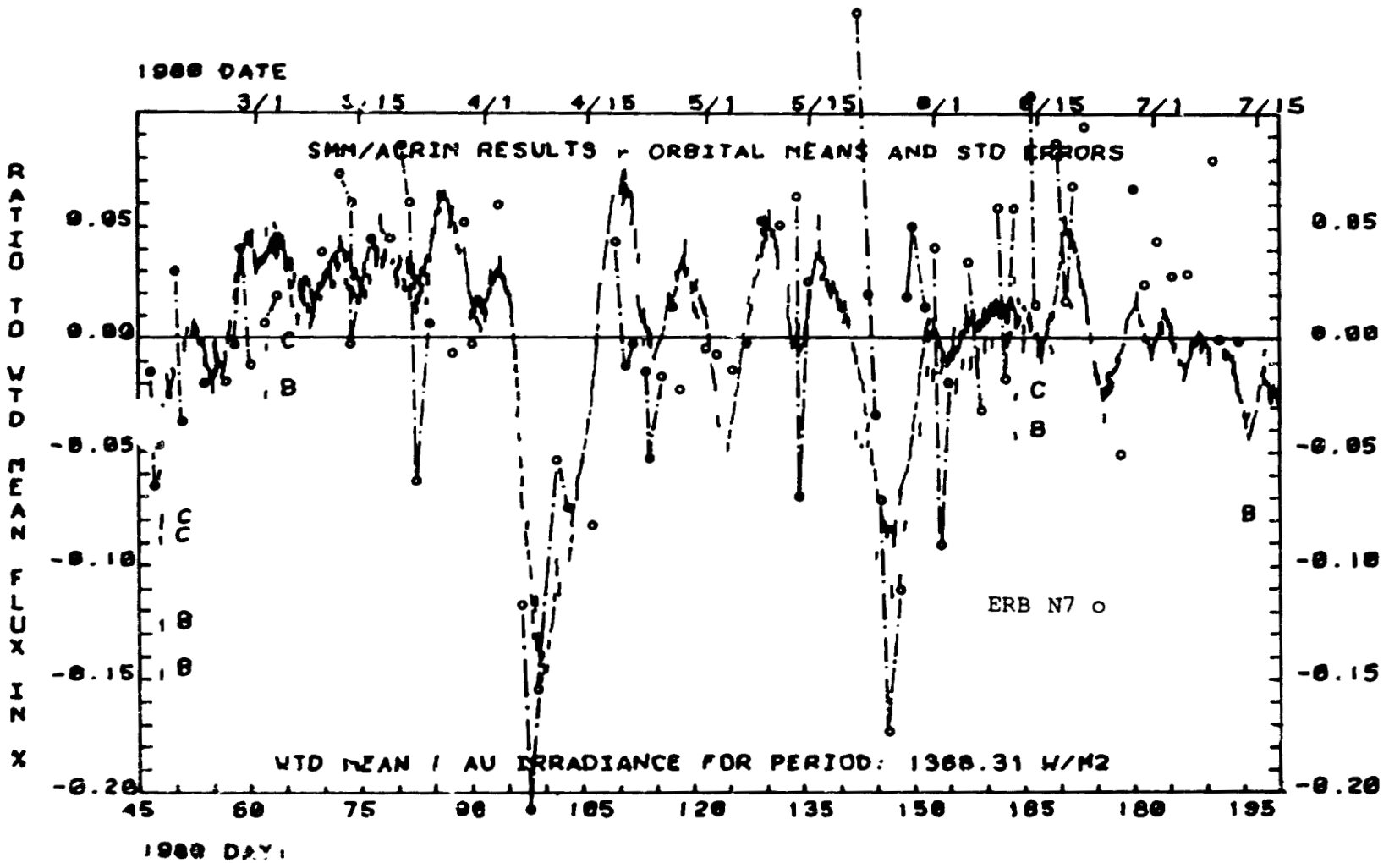


Figure 5. ERB channel 10c Engineering Data overlayed on SMM plot of Reference 5

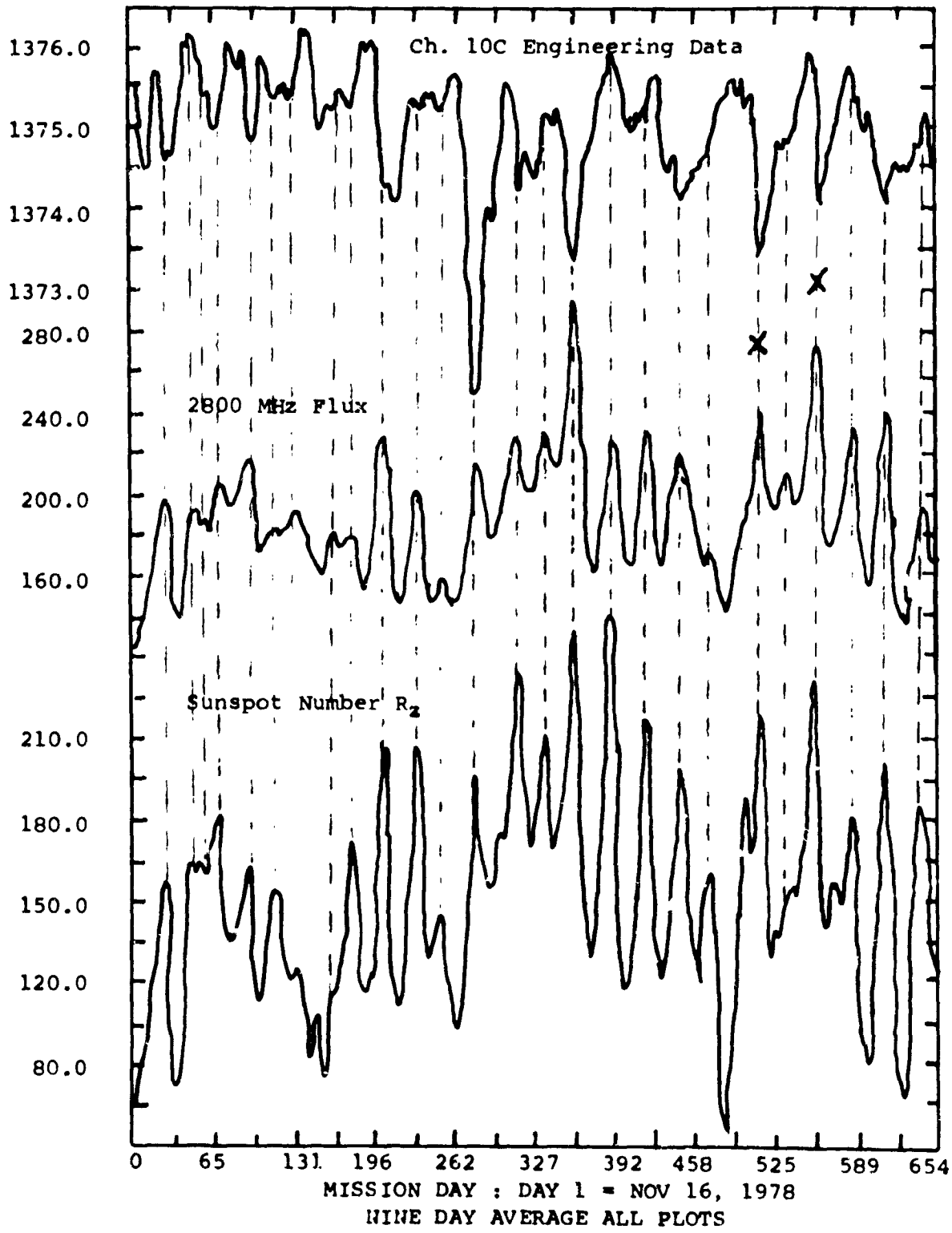


Figure 6. Plot of nine-day running means of 10c Engineering data and solar indicators for event recognition

INDIRECT METHODS FOR MEASURING VARIATIONS OF THE SOLAR CONSTANT

Sabatino Sofia
NASA/Goddard Space Flight Center

ABSTRACT

The various techniques thus far used to measure or infer variations of the solar constant, S, will be reviewed. The difference between the methods that measure δS, and those that measure variations in the solar luminosity, δL, is discussed. It is shown that the past practice of simply relating δS to δL by geometrical arguments is not valid because of anisotropy of the solar radiation. We conclude that direct techniques have proven the existence of short-term variability that is fully explainable in terms of the passage of active regions (spots and faculae) on the face of the Sun. These measurements, however, yield no conclusive evidence regarding variability on longer, climatically significant, time scales. The observations of changes in the solar diameter, on the other hand, support the existence of structurally induced variations of the solar luminosity on timescales of tens of years, which are clearly significant in our understanding of climatic variations.

INTRODUCTION

It has been common practice to use the concepts of solar luminosity, L, and solar constant or irradiance, S, interchangeably. L is the total energy output of the Sun per unit time, whereas S is the energy per unit time striking normally one unit area at the Earth's distance (and direction) from the Sun. If absorption can be neglected, and if the solar radiation is isotropic, these quantities are simply related,

$$L = 4\pi d^2 S, \qquad (1)$$

where d is the Sun-Earth distance*. Since until recently the simplifying assumptions required for equation (1) to be valid were unchallenged, the custom of referring to either L or S as equivalent concepts is understandable. We shall argue later on in this paper that the solar radiation is not isotropic. Since this invalidates relation (1), we shall hereafter clearly discriminate between S and L.

DIRECT METHODS

All direct methods for measuring the solar energy output involve determinations of S. This type of work was begun early in the century using ground-based detectors (ref. 1), and continued in the last two decades on the basis of observations carried out from various space-borne detectors (ref. 2). Of particular significance are the recent results from the ERB experiment on Nimbus 7 (ref. 3), and from the ACRIM experiment on the Solar Maximum Mission, SMM, (ref. 4) The importance of these two experiments, with overlapping measurements since February 1980, is that they both detected coincident variations of the solar constant, at levels of up to 0.2 percent, having timescales of days. These variations are now well explained in terms of active regions

* In practice, S is always normalized to a Sun-Earth distance of 1 astronomical unit.

(spots and faculae) brought into view and out of view by solar rotation (ref. 5).

On longer timescales, changes of S have not convincingly been measured (refs. 1, 2). First of all, calibration difficulties have detracted significance from any detected small changes. Second, since S does vary with a few days timescales, differences at the few tenths of a percent level obtained from different rocket or balloon flights might not reflect more than these active-region induced changes.

A crucial outcome of the Nimbus 7 and SMM observations, followed by its interpretation, is that the solar radiation is not isotropic, since the radiation emitted in a direction where a particularly large spot is visible, for example, is less than that emitted in any direction where the spot is not visible.

INDIRECT METHODS

Indirect methods have been used to determine variations of both S and L. The δS determinations from indirect techniques apply to geological and biological data, and have the properties of (1) are not very sensitive, and (2) apply to timescales in thousands of years, where no other (more precise) data exist. An example of this type of work is the ice-core probing to determine migrations of the ice-caps on the Earth's polar regions. We shall not discuss these measurements any further in this paper.

The indirect measuring techniques which are of interest to us measure changes of L on timescales smaller than a few hundred years. The common feature in all of them is that they are based on the Stefan-Boltzmann law

$$L = 4\pi\sigma R^2 T^4 \tag{2}$$

which gives the total energy output per unit time of a blackbody sphere of radius R and temperature T.

The Sun is not a blackbody, and equation (2) is not strictly applicable; however, if we substitute T by T_{eff}, the effective temperature, then equation (2) is valid. It must be stressed that the only means of obtaining T_{eff} is through relation (2), i.e $T_{eff} = (L/4\pi\sigma R^2)^{1/4}$, and so the relation is not useful to determine L. Let us take derivatives.

$$\delta L/L = 2\,\delta R/R + 4\,\delta T_{eff}/T_{eff} \tag{3}$$

We can now determine variations of the solar luminosity by measuring the radius changes, as well as the variations of the effective temperature, which may be a quantity much easier to measure than T_{eff} itself. Livingston (ref. 6) attempted to determine δT_{eff} by measuring spectroscopic temperature changes δT_s, determined from the strengths of weak Fraunhofer lines which orginate in the same layers of the solar atmosphere (the photosphere) as the bulk of the solar irradiance. Then

$$\frac{\delta T_s}{T_s} = \frac{\delta T_{eff}}{T_{eff}} \tag{4}$$

He chose for this purpose the line λ 5380.3 from CI. More recent work (ref.

7) indicates that the equivalent widths of different lines with diverse temperature sensitivities produce incompatible results. In fact, the current status of this particular problem requires a large amount of poorly understood modeling, and so for the time being we do not have a definitive way of measuring $\delta T_{eff}/T_{eff}$.

If one term of equation (3) cannot be determined, it would appear that the indirect techniques are useless to determine $\delta L/L$. We will show hereafter that such is not the case. In fact, changes of the solar structure (radius) are the cause of short-period luminosity changes, with δT_{eff} being but one of the consequences. It turns out, then, $\delta R/R$ can be directly related to $\delta L/L$ by means of numerical modeling. While our results are still model-dependent, they have advantages over the temperature model dependence in that (a) the modeling is better understood, and (b) it can be verified by observations.

THE RADIUS-LUMINOSITY RELATION

Let us assume that we have two stars of identical mass, age and initial chemical composition. Let us model them in the conventional manner (i.e. ignoring rotation, magnetic field, etc). Because of the nature of the equations of stellar structure and evolution, if we wish these stars to have a somewhat different luminosity, the only possible avenue is to use a different mixing length to model their envelope. This change will affect L, but also, and primarily, R. From this argument it was concluded that if the Sun is to acquire a different equilibrium value of its luminosity, it will do so at a different radius, and hence radius monitoring might be a sensitive means of monitoring changes of the solar luminosity (ref. 8). The trouble with this argument is that it cannot yield the relationship between R and L, since the timescales of interest do not always allow re-establishing total equilibrium (for example, the thermal timescale of the solar interior is $\gtrsim 10^6$ years), and the standard solar model does not contain all the physics that may lead to L changes on non-evolutionary timescales. By now extensive realistic numerical modeling of possible mechanisms leading to a quick luminosity change have been carried out (refs. 9, 10, 11). Contrary to our initial hope, it was found that the relationship between δL and δR is dependent on the mechanism that causes the changes, and more particularly, on the solar region where it primarily operates. If we define $W = \delta \ln R/\delta \ln L$, it is found that, in general, W is slowly time dependent. Three mechanisms have been examined to date, namely

1) α -mechanism: a sudden change of the mixing length on the solar convective envelope. This mechanism first proposed by Ulrich (ref. 12), has been examined extensively (refs. 9, 10, 11). The consensus is that $W \sim 6 \times 10^{-4}$, i.e. on short timescales the radius is very insensitive to $\delta\alpha$ which produce significant δL . The $\delta\alpha$, when applied to different regions of the convection zone, does not produce different results, since its effects are only important in the shallow super-adiabatic region (cf ref. 11).

2) β-mechanism : a layer at a given (variable) depth in the convection zone is perturbed by adding non-thermal pressure components (magnetic or turbulent pressure). Depending on the depth of the perturbation, this leads to a W that may reach $\sim .1$ (cf ref. 9).

3) core-perturbation: A sudden mixing event is arbitrarily induced within the partially nuclearly processed radiative core. In this case $W \leq .7$ (Sweigart, private communication).

Our current knowledge of W indicates that it depends strongly on the zone where the perturbation is effective, and to a lesser degree, on the size of the perturbation, and on the history and sequence of perturbing events (Endal, private communication). Further modeling is underway to better understand the sensitivity of W on the various parameters that may affect the Sun. It is now clear that the W value of the Sun, if indeed there is a unique value (say by the fact that one mechanism dominates the variations) cannot be determined from theory alone. Instead, we must determine W from simultaneous observations of δR and δL, after correcting δL for the modulation produced by active regions. The value of W thus obtained will allow us to determine δL for all times in which radius information is available, will identify the depth in the Sun where the changes originate, and finally, may identify the physical process responsible for the structural changes in the Sun.

CHANGES OF THE SOLAR RADIUS

It is convenient to separate measurements of the solar radius in two groups: those carried out in the past, and those currently underway or planned for the future. To date, three types of measurements carried out continually in the past have been identified which yield information on the solar radius. They are: (1) transit or meridian circle observations (refs. 13, 14), (2) timings of the transits of the planet Mercury in front of the Sun (refs. 15, 16), and (3) timings of total solar eclipses (refs. 16, 17). Because of the effect of the observer's personal equation (refs. 13, 14) and other unknown difficulties, the transit instrument timings cannot be literally taken as measurements of the solar radius. The transits of Mercury provide data apparently free of systematic errors, but the error of each individual measurement is $\sim \pm 1''$ (refs. 15, 16). Consequently, while they can successfully disprove large secular changes of the solar radius, they cannot provide any information regarding non-secular radius changes of amplitude $<$ few arc seconds. Finally, two types of solar eclipse timings have been proposed, namely (a) timing measurements carried out in the middle of the path of totality (ref. 16), and at the edge of the path of totality (ref. 17). Reasons have been stated (ref. 18) indicating that the path-edge observations, while fewer in number, provide the more reliable means of monitoring changes in the size of the Sun (ref. 19). The results of applying this technique thus far obtained are given in ref. 18. Of particular significance are the eclipses of 1925 and 1979, since they were observed by a large number of observers who were very near the edge of the path of totality (the timing error is negligible), whose location was extremely well documented (ref. 18), and three of the observed contacts occurred in the same lunar features, so that the derived radius change is independent of lunar profile errors. This shows that between 1925 and 1979 the solar radius differed by > 0.5.

Future measurements should include edge timings to link all future results to past results. In addition, however, measurements by more modern techniques will be carried out. In particular, the SCLERA telescope should be able to detect radius changes at the milli-arc second level, a factor of 50 better than the edge timing observations (ref. 20). Additional instrumentation is currently planned at the High Altitude Observatory and at Sacramento Peak Observatory.

IMPLICATIONS OF THE SOLAR RADIUS CHANGES

Timing observations near the edge of the path of totality have shown that the solar radius changes by approximately 0.5 in timescale of tens to hundreds of years. This implies that solar luminosity variations of structural origin having similar timescales have taken place. Notice that these changes would, in addition, undergo the activity-induced modulation described in ref. 5. In fact, the occurrence of structurally induced changes will show up as a secular growth in the residuals of the analysis given in ref. 5. To estimate what 0.5 change corresponds to in terms of solar luminosity change, we must know the mechanism responsible for the variation. The occurrence of such large radius changes already eliminates the α-mechansim, since it would imply a $\delta L/L > 80$ percent, a value clearly excluded by the history of the Earth's climate. If the change originated by events occurring in the solar core, the implied luminosity change would be at the 0.03 percent level, thus negligible for most climatic purposes. However, it is difficult to visualize interior mechanisms acting on timescales as short as tens of years. A more probable origin of the solar variations resides in a β-mechanism associated with the variable magnetic field produced at the base of the convection zone by the solar dynamo. In this case, $\delta L/L$ may be of the order of 0.5 percent, a value neither excluded by the climatic history of the Earth, nor negligible. In fact, the size of the climatic effect would strongly depend on the duration of the changes, and this cannot be currently assessed on the basis of only 5 radius values with rather uneven time-distribution (ref. 18).

SUMMARY AND DISCUSSION

From direct measurements carried out by detectors on board NIMBUS 7 and SMM, it has been found that the solar irradiance varies with timescales of days to months. Such variations can be fully explained in terms of flux modulation caused by the passage of active regions on the visible solar hemisphere, and reflect the directional dependence of the energy deficit caused by the sunspots and the re-emission of the energy by the faculae. The measurements do not require (or support) variations of the solar luminosity during the six months of high-precision observations carried out by the SMM.

Indirectly, solar luminosity changes with timescales of tens to hundreds of years can be inferred from structural changes in the Sun as revealed from past radius measurements. While the factor relating radius changes to luminosity changes, which depends on the mechanism that produces the variations, is not known with complete confidence, the currently favored interpretation (a β-mechanism) supports luminosity variations having an amplitude of $< 0.5\%$. These variations, if long-lived, have important climatic consequences.

REFERENCES

1. Angione, R. J.: Review of Ground-Based Measurements. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

2. Fröhlich, C.: The Variability of the Solar Output. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

3. Hickey, J. R., B. M. Alton, F. J. Griffin, H. Jacobowitz, P. Pellegrino, E. A. Smith, T. H. Vonder Haar, and R. H. Mashoff: Solar Variability Indications from Nimbus 7 Satellite Data. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

4. Willson, R. C.,S. Gulkis, M. Jannsen, H. S. Hudson, and G. A. Chapman: Observations of Solar Irradiance Variability. Science, 211, 700, 1981.

5. Oster, L. F., K. H. Schatten, and S. Sofia: Solar Irradiance Variations due to Active Regions. Astrophys. J. Letters (submitted for publication), 1981.

6. Livingston, W. C.: Cooling of the Sun's Photosphere Coincident with Increased Sunspot Activity. Nature, 272, 340, 1978.

7. Livingston, W. C., H. Holweger and O. R. White: Observed Variability in the Fraunhofer Line Spectrum of Solar Flux, 1975-1980. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

8. Sofia, S., J. O'Keefe, J. R. Lesh, and A. S. Endal: Solar Constant: Constraints on Possible Variations Derived from Solar Diameter Measurements. Science, 204, 1306, 1979.

9. Chan, K. L., and S. Sofia: Estimating Short Term Solar Variations by a Simple Envelope Matching Technique. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

10. Twigg, L. W., and A. S. Endal: Thermal Perturbation of the Sun. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

11. Sweigar , A. V.: Effects of Changes in Convective Efficiency on the Solar Radius and Luminosity. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

12. Ulrich, R. K.: Solar Neutrinos and Variations in the Solar Luminosity. Science, 190, 619, 1975.

13. Eddy, J. A.: Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

14. Smith, C., and D. Messina: The Horizontal and Vertical Semi-Diameters of the Sun Observed at the Cape of Good Hope (1834-1887) and Paris (1837-1906): a Report on Work in Progress. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

15. Shapiro, I. I.: Is the Sun Shrinking? Science, 208, 51, 1980.

16. Parkinson, J. H., L. V. Morrison and F. R. Stephenson: The Constancy of the Solar Diameter Over the past 250 years. Nature, 288, 548, 1980.

17. Sofia, S., D. W. Dunham and A. D. Fiala: Determination of Variations of the Solar Radius from Solar Eclipse Observations. Proc. Conf. Ancient Sun, (Pepin, R., J. A. Eddy, and R. B. Miller, editors) pp 147-157, 1980.

18. Dunham, D. W., J. B. Dunham, A. D. Fiala and S. Sofia: Eclipse Radius Measurements. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

19. Dunham, D. W., S. Sofia, A. D. Fiala, D. Herald, and P. M. Muller: Observations of a Probable Change in the Solar Radius Between 1715 and 1979. Science, 210, 1243, 1980.

20. Hill, H., P. Caudell, and R. Bos: Workshop on Variations of Solar Constant, NASA CP-2191, 1981.

SOLAR NEUTRINOS*

John N. Bahcall
Institute for Advanced Study
Princeton, NJ 08540

INTRODUCTION

The topics to be covered are, in order: an overview of the subject of solar neutrinos, a brief summary of the theory of stellar evolution, a description of the main sources of solar neutrinos, a brief summary of the results of the Brookhaven ^{37}Cl experiment, an analysis of the principal new solar neutrino experiments that have been proposed, a discussion of how solar neutrino experiments can be used to detect the collapse of stars in the Galaxy, and finally, a description of how the proposed ^{71}Ga experiment can be used to decide whether the origin of the present discrepancy between theory and observation lies in our conventional solar models or our conventional physics.

The most important fact is that there is a serious discrepancy between the standard theory and observation.

Neutrinos can be used directly and quantitatively to test the theory of nuclear energy generation in stars like the sun. Of the particles released by the assumed thermonuclear reactions in the solar interior, only neutrinos have the ability to penetrate from the center of the sun to the surface and escape into space. Thus neutrinos offer us a unique possibility of "looking" into the solar interior.

The theory of stellar aging by thermonuclear burning is widely used in interpreting many kinds of astronomical information and is a necessary link in establishing such basic data as the ages of the stars and the abundances of the elements. The parameters of the sun (its age, mass, luminosity, and chemical composition) are better known than those of any other star, and it is in the simplest and best understood stage of stellar evolution, the quiescent main sequence stage. Thus an experiment designed to capture neutrinos produced by solar thermonuclear reactions is a crucial one for the theory of stellar evolution.

A number of exotic solutions to the solar neutrino problem, modifying either the physics or the astronomy (and in some cases both), have been proposed. Even if one grants that the source of the discrepancy is astronomical, there is no general agreement as to what aspect of the theory is most likely to be incorrect. As indicated above, many of the proposed solutions of the solar neutrino problem have broad implications for conventional astronomy and cosmology. Some of them would change the theoretical ages of old stars or the inferred primordial element abundances. On the other hand, modified theories of the weak interactions have been proposed in which neutrinos may disappear

*This work was supported by the NSF Contract PHY79-19884.

by mixing or decay in transit from the sun to the earth, but for which there are no terrestrially measurable consequences. It is conceivable that one of these modified theories of the weak interactions is correct and the standard solar model is not in conflict with observations.

STELLAR EVOLUTION

Table 1 summarizes the principles that are required for constructing solar models that are tested by solar neutrino experiments. There are many more things in stellar evolution theory, but they are not essential for understanding solar neutrino experiments, certainly not for the purposes of this talk.

The first principle is hydrostatic equilibrium, which in practice is used together with the special assumption of spherical symmetry. The second principle is that the energy source is postulated to be nuclear; the rates of the nuclear reactions depend on the density (ρ) and the temperature (T), and the composition (X_i). The practical part of this principle is that the rate at which the nuclear reactions produce energy when integrated over the whole sun is equal to the observed solar luminosity today. The "today" is an essential part of this principle.

The third principle is that the energy is transported from the deep interior to the surface via radiation and convection. In practice, for most (but not quite all) of the models, the great bulk of the energy is transported by radiation. The key quantities are the gradient of the temperature (dT/dr) and the opacity of the solar matter.

The assumption that the initial composition was uniform and is equal to the presently observed surface composition is closely related to the question of which opacity should be used. It is plausible that the surface composition has not changed much because of nuclear reactions since the sun was formed. It is not quite so obvious that nothing has been added to the solar surface since the sun was born. However, that is the assumption which is widely used throughout astronomy and is the basis for making the standard calculations.

The final principle is that the sun evolves because it burns its nuclear fuel. It has burned for something like 5 billion years so far. One mocks up this evolution by computing several quasistatic models which march along in time.

The bottom line of this brief course in stellar evolution is: within our store of observational information about stars, only the Brookhaven Chlorine 37 experiment of Ray Davis and his colleagues is inconsistent with the standard theory of stellar evolution. It is the only place where we do not see a result of observational difficulties unless we modify something among the basic assumptions.

NUCLEAR FUSION IN THE SUN

I shall now outline briefly the conventional wisdom regarding nuclear fusion as the energy source for main sequence stars like the sun. It is assumed that the sun shines because of fusion reactions similar to those envisioned for terrestrial fusion reactors. The basic solar process is the fusion of four protons to form an alpha particle, two positions (e^+), and two neutrinos (ν), that is, $4p \rightarrow \alpha + 2e^+ + 2\nu_e$. The principal reactions are shown in Table 2 with a column indicating in what percentage of the solar terminations of the proton-proton chain each reaction occurs. The rate for the initiating proton-proton (p-p) reaction, number 1 in Table 2, is largely determined by the total luminosity of the sun. Unfortunately, these neutrinos are below the threshold, which is 0.81 MeV, for the ^{37}Cl experiment. Several of the proposed new experiments, especially the ^{71}Ga experiment, will be primarily sensitive to neutrinos for the p-p reaction. The p-e-p reaction (number 2), which is the same as the familiar p-p reaction except for having the electron in the initial state, is detectable in the ^{37}Cl experiment. The ratio of p-e-p to p-p neutrinos is approximately independent of which model (see below) one uses for the solar properties. Two other reactions in Table 2 are of special interest. The capture of electrons by ^{7}Be (reaction 6) produces detectable neutrinos in the ^{37}Cl experiment. The ^{8}B beta decay, reaction 9, was expected to be the main source of neutrinos for the ^{37}Cl experiment because of their relatively high energy (14 MeV), although it is a rare reaction in the sun (see Table 2). There are also some less important neutrino-producing reactions from the carbon-nitrogen-oxygen (CNO) cycle, but we shall not discuss them in detail since the CNO cycle is believed to play a rather small role in the energy-production budget of the sun.

THE BROOKHAVEN SOLAR NEUTRINO EXPERIMENT

The Brookhaven solar neutrino detector is based on the neutrino capture reaction (refs. 1,2)

$$\nu + {}^{37}Cl \underset{\text{decay}}{\overset{\text{capture}}{\rightleftharpoons}} {}^{37}Ar + e^- \qquad (1)$$

which is the inverse of the electron capture decay of ^{37}Ar. The radioactive decay occurs with a half-life of 35 days. This reaction was chosen for the Brookhaven solar neutrino experiment because of its unique combination of physical and chemical characteristics, which were favorable for building a large-scale solar neutrino detector. Neutrino capture to form ^{37}Ar in the ground state has relatively low energy threshold (0.81 MeV) and a favorable cross section, nuclear properties that are important for observing neutrinos from ^{7}Be, ^{13}N, and ^{15}O decay and the p-e-p reaction.

The ^{37}Cl reaction is very favorable from a chemical point of view. Chlorine is abundant and inexpensive enough that one can afford the many hundreds of tons needed to observe solar neutrinos. The most suitable chemical compound is perchloroethylene, C_2Cl_4, a pure liquid, which is manufactured on a large scale for cleaning clothes. The product, ^{37}Ar, is a noble gas, which should ultimately exist in the liquid as dissolved atoms. The neutrino capture process produces an ^{37}Ar atom with sufficient recoil energy to break free

of the parent perchlorethylene molecule and penetrate the surrounding liquid, where it reaches thermal equilibrium.

The Brookhaven ^{37}Cl detector was built by Davis deep underground to avoid the production of ^{37}Ar in the detector by cosmic rays. This was done with the cooperation of the Homestake Gold Mining Company (Lead, South Dakota), who excavated a large cavity in their mine (~ 1500 m below the surface) to house the experiment. The final detector system consists of an ~400,000 liter tank of perchloroethylene, a pair of pumps to circulate helium through the liquid, and a small building to house the extraction equipment.

A set of 39 experimental runs carried out in the Brookhaven ^{37}Cl experiment over the last 10 years show that the ^{37}Ar production rate in the tank is 0.50 ± 0.06 ^{37}Ar atoms per day (refs. 1,2,3). Even though the tank is nearly a mile underground, a small amount of ^{37}Ar is produced by cosmic rays An evaluation of data obtained by exposing 7500 liters of C_2Cl_4 at various depths underground suggests that the cosmic-ray production rate in the detector may be 0.08 ± 0.03 ^{37}Ar atoms per day (refs. 2,3), Fireman's (ref. 4) measurements of the muon background using a ^{37}K detector suggest a background rate of (0.18 ± 0.09) ^{37}Ar atoms per day. If this background rate is correct, then there is no evidence for any solar neutrino detection beyond the 3-σ level of significance.

If the background rate determined from the C_2Cl_4 measurements is assumed, then a positive signal of (2.2 ± 0.4) SNU is inferred (refs. 1,2,3) (1SNU = 10^{-36} captures per target particle per second).

The predicted capture rates for one recently constructed solar model are shown in Table 3 (ref. 5). The results are expressed in terms of SNU's = 10^{-36} captures per target atom per second, the characteristic counting rate for solar neutrino experiments. The neutrino absorption cross sections used to compute the rates given in Table 3 are from reference 6. The best values to use for various nuclear parameters is currently under investigation and the total predicted rate may well differ by as much as 1 to 1.5 SNU from the value of 7.8 SNU shown in Table 3 (and be as low as 5 SNU if the preliminary cross-section measurement of $^{3}He(\alpha,\gamma)^{7}Be$ by C. Rolf *et al.* is verified).

OBSERVATIONAL IMPLICATIONS

The ^{37}Cl experiment tests theoretical ideas at different levels of meaning, depending on the counting rate being discussed. The various counting rates and their significance are summarized in Table 4. It is obvious from a comparison of Table 4 with the experimental results given above that the value of 28 SNU's based on the CNO cycle is ruled out. More surprisingly, the best current models based on standard theory, which imply ~6 to 8 SNU's are also inconsistent with the observations. This disagreement between standard theory and observation has led to many speculative suggestions of what might be wrong. One such suggestion (ref. 7), that in the solar interior the heavy element abundance is at least a factor of 10 less than the observed surface abundance, leads to an expected counting rate of 1.5 SNU's (see Table 3), which is about as low a prediction as one can obtain from solar models without seriously changing current ideas about the physics of the solar

interior. We note that present and future versions of the ^{37}Cl experiment are not likely to reach a sensitivity as lc as 0.3 SNU, the minimum counting rat. (from reaction 2 of Table 2) that can be expected if the basic idea of nuclear fusion as the energy source for main sequence stars is correct.

NEW EXPERIMENTS

Another experiment is required to settle the issue of whether our astronomy or our physics is at fault. Fortunately, one can make a testable distinction. The flux of low energy neutrinos from the p-p and p-e-p reactions (numbers 1 and 2 in Table 2) is almost entirely independent of astronomical uncertainties and can be calculated from the observed solar luminosity, provided only that the basic physical ideas of nuclear fusion as the energy source for the sun and of stable neutrinos are correct. If these low energy solar neutrinos are detected in a future experiment, we will know that the present crisis is caused by a lack of astronomical understanding. If the low energy neutrinos are absent, we will know that the present discrepancy between theory and observation is due at least in part to faulty physics, not just poorly understood astrophysics.

I have analyzed in detail the theoretical aspects of eleven experiments that have been studied by various experimental groups as possible new solar neutrino experiments (ref. 6). Those eleven proposed targets are listed in the first column of Table 5. I also list in the other columns the following information: (a) whether the total cross-section solar neutrinos can be calculated to an accuracy of at least ten percent; (b) whether something new will be learned about the solar interior, or neutrino physics, by performing the proposed experiment; and (c) whether (in my opinion) the experiment is feasible with current technology. A check mark (√) indicates that the answer to the relevant question is affirmative; a negative answer is indicated by (X).

The detectors for solar neutrinos can be classified according to their relative sensitivity to different parts of the solar neutrino spectrum. Five of the experiments are primarily sensitive to ^{8}B neutrinos; these are ^{2}H, ^{37}Cl, ^{81}V, ^{55}Mn, and neutrino-electron scattering.

Four detectors, ^{71}Ga, ^{87}Rb, ^{115}In, and ^{205}Tl, are primarily sensitive to neutrinos from the proton-proton reactions. The expected capture rates for these detectors are practically independent of the astronomical assumptions that are made provided only that the sun produces, in a steady-state fashion and via the proton-proton chain, the energy that it radiates from its surface.

The p-e-p neutrinos (reaction 2, Table 2) are expected to make the largest single contribution to the capture rate of a ^{7}Li detector, even for the standard solar model. The observational results from the ^{37}Cl experiment show, moreover, that the higher energy ^{8}B neutrinos should contribute, for a ^{7}Li target, at most one-half the capture rate due to p-e-p neutrinos. Since the p-e-p neutrinos are as good a measure of the proton-proton reaction rate as are the p-p neutrinos, one can also classify the ^{7}Li detector as a p-p sensitive target. The ^{7}Li and ^{115}In targets share the property of being reasonably sensitive to more than one neutrino branch (the p-e-p, ^{7}Be, ^{8}B, and ^{15}O branches

for the ^{7}Li detector; the p-p and ^{7}Be branches for the ^{115}In target). The p-p and ^{7}Be capture rates could be determined separately for the ^{115}In experiment since the energies of the individual electrons could be measured.

The ^{81}Br detector is primarily sensitive to ^{7}Be neutrinos.

The ^{115}In and neutrino-electron scattering experiments could in principle be used to measure the direction of the electrons that are produced and thus to establish that the incident neutrinos come from the sun.

In order for a solar neutrino experiment to be most useful, the absorption cross sections must be accurately known. Of the new targets discussed in this paper, only ^{2}H, ^{7}Li, ^{71}Ga, ^{87}Rb, ^{115}In, and neutrino-electron scattering satisfy this requirement. A new detector should also help discriminate between the possible explanations of the discrepancy between theory and observation in the ^{37}Cl experiment. Experiments with ^{2}H or neutrino-electron scattering are sensitive primarily to ^{8}B neutrinos, as is the ^{37}Cl experiment. In order to provide new information of astrophysical importance, these experiments must be sensitive to a ^{8}B flux that is significantly below that already reached by the Brookhaven ^{37}Cl experiment. There has not been a recent and detailed experimental feasibility study for the proposed ^{87}Rb experiment, perhaps because of the uncomfortably short lifetime (2.8^{hrs}) of the daughter nucleus, ^{87}Sr. If we set aside ^{87}Rb because of the absence of a feasibility study, then the preferred targets are: ^{7}Li, ^{71}Ga, ^{115}In, and either ^{2}H or electron scattering (if sufficiently sensitive).

There are four major neutrino branches that must be measured in order to carry out a program of neutrino spectroscopy of the solar interior. These branches are the p-p, ^{7}Be, ^{8}B, and $^{13}N + ^{15}C$ neutrinos. The future experimental solar neutrino program should include all of the preferred new detectors. The ^{71}Ga experiment is primarily sensitive to p-p neutrinos and the ^{37}Cl experiment to ^{8}B neutrinos. The ^{7}Li and ^{115}In experiments provide additional information about the ^{7}Be and $^{13}N + ^{15}O$ fluxes. Taken together, the results of the four experiments (^{7}Li, ^{37}Cl, ^{71}Ga, and ^{115}In) should allow us to solve for the parameters of the solar interior (temperature range density and composition). An ^{2}H or an electron-neutrino experiment should also be performed at some future date in order to check on the upper limit to the ^{8}B flux determined by the ^{37}Cl experiment. If a feasible experiment is proposed in which a ^{8}B flux as low as twenty percent of the prediction from the standard model could be measured then this would also be a preferred experiment since it would provide qualitatively new astrophysical information.

Either a ^{71}Ga or an ^{115}In experiment can distinguish between explanations that are based on presumed inadequacies in, respectively, the astronomical theory or the weak interaction theory provided only that the sun produces in a steady-state fashion the energy it radiates from its surface. A low counting rate in either of these experiments could also arise, in principle, if the sun is now in an abnormal phase in which its nuclear energy generation is much less than its surface luminosity. However, for most of the models of this kind that have appeared in the literature, the reduction in the counting rate of a ^{71}Ga or an ^{115}In experiment would not be nearly as great as is expected on either the oscillation or the decay hypothesis. Moreover, these latter two

processes lead to specific predictions for the ^{71}Ga and ^{115}In experiments when combined with the results of the ^{37}Cl experiment.

THE ^{71}Ga EXPERIMENT

A preliminary background experiment has been completed with 1.3 tons of gallium and plans are underway for an approximately 10-ton calibration experiment to be performed in a mine in Tennessee using a strong (megacurie) radioactive source made in the reactor at Oak Ridge. This program involves an international collaboration (ref. 8) between Brookhaven National Laboratory, the Max-Planck Institute for Nuclear Physics at Heidelberg, the Institute for Advanced Study, and the Weizmann Institute.

STELLAR COLLAPSES

It is now generally believed by workers in the field that much of the potential energy which is released when stars collapse is emitted in the form of neutrinos (see, *e.g.*, refs. 9 and 10). The rate at which optically undetected stellar collapses occur in the galaxy is not known, but plausible estimates might vary from once a year to once in a hundred years. The solar neutrino experiments that have been discussed in this talk can all be used to detect occasional stellar collapses in the galaxy. In fact, I have even suggested that the exceptionally high result observed in Run 27 of the ^{37}Cl experiment might be interpreted in terms of a stellar collapse.

Using plausible parameters for the energy liberated in the form of neutrinos, I estimate (ref. 6) that stellar collapses can be detected to a typical distance of a few kpc with the proposed solar neutrino detectors. The specific values are: ^{7}Li (3 kpc), ^{37}Cl (4.5 kpc), ^{71}Ga (≥ 0.3 kpc), ^{115}In ($\gtrsim 10$ kpc), ^{2}H (~ 10 kpc), and neutrino-electron scattering experiments ($\lesssim 2$ kpc).

C - 2

REFERENCES

1. Davis, R.: Proceedings of the Brookhaven Solar Neutrino Conference, vol. 1, 1978, p. 1.

2. Bahcall, J.N. and Davis, R., Jr.: Science, vol. 191, 1976, p. 246.

3. Rowley, J.K., Cleveland, B.T., Davis, R., Jr., Hampel, W., Kirsten, T.: The Present and Past Neutrino Luminosity of the Sun. BNL preprint, no. 27190, 1980.

4. Fireman, E.L.: 16th International Cosmic Ray Conference, vol. 12, 1979. Kyoto, Japan.

5. Bahcall, J.N., Lubow, S.H., Heubner, W.F., Magee, N.H., Merts, A.L., Argo, M.F., Parker, P.D., Rozsnyai, B., and Ulrick, R.K.: Phys. Rev. Lett., vol. 45, Sept. 15, 1980, pp. 945-948.

6. Bahcall, J.N.: Rev. Mod. Phys., vol. 50, 1978, p. 881.

7. Bahcall, J.N. and Ulrich, R.K.: Ap. J., vol 170, 1971, p. 593.

8. Bahcall, J.N., Cleveland, B., Davis, R., Jr., Dostrovsky, I., Evans, J.C., Jr., Fratti, W., Friedlander, G., Lande, K., Rowley, K., Stoner, D., and Weneser, J.: Phys. Rev. Lett., vol. 40, 1978, p. 1351.

9. Colgate, S. and White, R: Ap. J., vol. 143, 1966, p. 626.

10. Brown, G.E.: Comments on Astrophysics, vol 7, 1977, p. 67.

Table 1. Three-minute Course in Stellar Evolution Principle

Hydrostatic Equilibrium
Spherical Sun
Nuclear Energy Source
Energy Transport by Radiation & Convection
Uniform Primordial Composition = Surface Composition
Evolution (age = 5×10^9 yrs.)
BOTTOM LINE: Only ^{37}Cl Experiment Inconsistent with Standard Theory

Table 2. The Proton-Proton Chain in the Sun

Number	Reaction	Solar terminations (%)	Maximum Neutrino Energy (Mev)
1	$p+p \rightarrow {}^2H+e^+ +\nu$	(99.75)	0.420
	or		
2	$p+e^- +p \rightarrow {}^2H+\nu$	(0.25)	1.44 (monoenergetic)
3	${}^2H+p \rightarrow {}^3He+\nu$	(100)	
4	${}^3He+{}^3He \rightarrow {}^4He+2p$	(86)	
	or		
5	${}^3He+{}^4He \rightarrow {}^7Be+\nu$	(14)	
6	${}^7Be+e^- \rightarrow {}^7Li+\nu$		0.86 (90%), 0.383 (10%) (Both monoenergetic)
7	${}^7Li+p \rightarrow 2{}^4He$		
	or		
8	${}^7Be+p \rightarrow {}^8B+\nu$	(0.02)	
9	${}^8B \rightarrow {}^8Be^* +e^+ +\nu$		14.06
10	${}^8Be^* \rightarrow 2{}^4He$		

Table 3. Predicted Capture Rates for a Recently Computed Solar Model[5]

Neutrino Source	Capture Rate (SNU's)
p-p	0
^{8}B	6.3
PEP	0.2
^{7}Be	1
^{13}N	0.06
^{15}O	0.2
Total = 7.8 SNU	

Table 4. Significance of Counting Rates in the ^{37}Cl Experiment. One Solar Neutrino Unit (SNU) = 10^{-36} Captures per Target per Second.

Counting Rate (SNU)	Significance of counting rate
28	Expected if the CNO cycle produces the solar luminosity
7±1	Predictions of current models
1.5	Expected as a lower limit consistent with standard ideas of stellar evolution
0.3	Expected from the PEP reaction, hence a test of the basic idea of nuclear fusion as the energy source for main sequence stars

Table 5. Proposed Experiments

Target	Cross-Section	New	Feasible
^{2}H	✓	?*	✓
^{7}Li	✓	✓	✓(?)
^{37}Cl	✓	✓	✓
^{51}V	X	X	✓
^{55}Mn	X	X	✓
^{71}Ga	✓	✓	✓
^{81}Br	X	✓	✓
^{87}Rb	✓	✓	X(?)
^{115}In	✓(?)	✓	✓(?)
^{205}Tl	X	✓	✓(?)
ν_e-$\bar{e}$	✓(W-S)	?*	?

*New if ϕ (^{8}B) ~ 6×10^{5} cm^{-2} sec^{-1} measurable

OBSERVED VARIABILITY IN THE FRAUNHOFER LINE SPECTRUM OF SOLAR FLUX, 1975-1980

W. Livingston
Kitt Peak National Observatory*

H. Holweger
Institut für Theoretische Physik
und Sternwarte, Kiel, West Germany

O. R. White
High Altitude Observatory
National Center for Atmospheric Research

ABSTRACT

Over the past five years double-pass spectrometer observations of the "sun-as-a-star" have revealed significant changes in line intensities. The photospheric component has weakened linearly with time 0 to 2.3%. From a lack of correlation between these line weakenings and solar activity indicators like sunspots and plage we infer a global variation of surface properties. Model-atmosphere analysis suggests a slight reduction in the lower-photospheric temperature gradient corresponding to a 15% increase in the mixing length within the granulation layer. Chromospheric lines such as Ca II H and K, Ca II 8542 and the CN band head weaken synchronously with solar activity. Thus the behavior of photospheric and chromospheric lines is markedly different, with the possibility of secular change for the former.

INTRODUCTION

Since 1975/1976 systematic observations have been routinely made at Kitt Peak in an attempt to quantify any temporal variation in the Fraunhofer lines. The original hope was that temperature sensitive lines might provide an indication of luminosity variation, assuming $F = \sigma T^4$ (ref. 1). When line strengths are measured near disk center variations of several percent are the rule, both in time and space, owing to a hierarchy of surface inhomogeneities—the granulation, supergranulation, faculae, plage, and so on. But if we observe the whole disk at once, in "integrated light", this fine structure averages out and the variance in line equivalent widths between spectrum scans reduces to

* Operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

~0.02%. This very low noise arises because the equivalent width parameter refers to the integral of absorption across the entire line profile, and so is invariant to instrument resolution, and because equivalent width is taken relative to the local continuum and thus is invariant to the transmission fluctuations of the earth's atmosphere. In terms of the predicted change in the high excitation line of CI 5380 our sensitivity is 1 K, or 0.1% in the solar constant assuming the sun is a black body.

After 2 years of data were acquired we noted that CI 5380 was 0.6% weaker than at the beginning and, following the above, we concluded the sun had cooled 6 K (ref. 2). At that point the equivalent width values for the other lines were really inconclusive. Today, however, with 5 years of data before us, the simple picture of temperature change is no longer tenable (ref. 3). As will be shown below all photospheric lines have either weakened or remain fixed, and the degree of weakening obviously depends on other factors besides temperature.

In the sections which follow, the full data set for both photospheric and chromospheric lines is presented. A tentative explanation of photospheric line behavior emerges from a full atmospheric-model analysis. The interpretation of the chromospheric Ca II H and K lines appears straightforward in terms of plage.

PHOTOSPHERIC LINES

OBSERVATIONS

All data used in this study were taken with the 13.5 m double-pass grating spectrometer of the McMath Telescope on Kitt Peak. Details of the physical instrument and reduction procedures are given elsewhere (refs. 4, 5, 6).

Table I lists the observed lines together with pertinent line parameters such as excitation potential, Zeeman sensitivity, and temperature sensitivity. Figure 1 gives examples of the run of equivalent width with time. Within the noise band a linear fit adequately represents the data up to the present. The 5 year intercept of this fit, converted to fractional change, is tabulated in the last column of table I.

What is the cause of the scatter in equivalent widths displayed in figure 1? If the scatter were solar in origin, say caused by facular line weakening which at high resolution was discovered by Chapman and Sheeley (ref. 7) then the weakening of different lines should correlate on a day-to-day basis. None is found (ref. 8). We conclude this scatter is instrumental in origin, perhaps arising from several causes such as spectrograph misalignment, grating drive screw error and atmospheric scintillation. Because these errors are all stochastic in nature, with no known secular trend, we accept the scatter as unavoidable with the present technique, but find no reason to doubt the long

term changes. Note that the equivalent width of Si I 10827 has remained constant (fig. 1).

INTERPRETATION

In figure 2 we explore a number of mechanisms to explain the obs[illegible] tions, (ref. 8). First we evoke a step change in temperature through [illegible] line-forming region but immediately find a conflict in sign between ca[illegible] the other lines (fig. 2a). Microturbulance can affect the strength of wea[illegible] lines, but proves ineffectual here (fig. 2b). Likewise Zeeman broadening alters line strength but doesn't work either (fig. 2c). The best fit between theory and observations follows from a 2.3% change in surface chemical abundance (fig. 2d) but this is considered unacceptable in principle.

Having disposed of the more simple possibilities we turn to the fact that our various lines are formed at differing levels in an atmosphere which possesses a temperature gradient. As shown by the respective contribution functions, C 5380 is formed low, Fe 5250 high, while Si 10827 is common to both regions (fig. 3). The positive thermal response of the carbon line can be combined with the negative response of the iron lines (table I) if we introduce a cooling of the low photosphere while heating up the higher layers. This can be accomplished by a change in the ratio of mixing length to pressure scale height ℓ/H from 2 to 2.3 (fig. 4). This initial value of $\ell/H=2$ is in accord with the recent evolutionary calculations for the sun by Mazzitelli. If we assume this 15% change in ℓ/H is confined to the granulation layer, the luminosity effect is negligible (ref. 9).

CHROMOSPHERIC LINES

OBSERVATIONS

Figure 5 displays the essence of the variation seen in the calcium K profile, which occurs in the line core primarily, if not wholly, and arises from the contribution of plage over the visible disk. As described elsewhere (ref. 6), we measure both the central intensity of the line (K_3 intensity) but also, for comparison with stellar observers, the 1 A equivalent width (K-index). The temporal behavior of K_3-intensity and K-index are essentially the same, figure 6, except the modulation of the K-index somewhat less than the former. The proof that the Ca K-line variation is due to chromospheric plage is given by the correlation of K_3-intensity with Mt. Wilson plage area index, figure 7. Similar results are found for Ca II 8542, and CN 3883, although the correlation coefficient is less (ref. 10).

CONCLUDING REMARKS

We have seen that over our 5 year time base photospheric lines have weakened linearly with time while chromospheric indicators mimic the activity cycle. Remaining unanswered are the following questions:

1. Is the photospheric variation truly global, or does it reflect unresol ed elements of solar activity (faculae, for instance)?

2. Is the time scale of the photospheric variation tied to the (22 y) solar cycle or is it evidence of some secular change of unknown duration?

3. In effect we are observing a change in line-blanketing. Will this affect solar irradiance?

REFERENCES

1. Livingston, W.; Milkey, R.; and Slaughter, C.: Solar Luminosity Variation I. CI 5380 as a Temperature Indicator and a Search for Global Oscillations. Ap. J. 211, 1977, pp. 281-287.

2. Livingston, W. C.: Cooling of the Sun's Photosphere Coincident with Increased Sunspot Activity. Nat. 272, 1978, pp. 340-341.

3. Hill, Henry A.; Livingston, William C.; and Caudell, Thomas, P.: Empirical Sensitivity of the CI 5380 Line to Temperature Changes in Static and Oscillating Systems. Ap. J. 214, 1977, pp. L137-L140.

4. Brault, J. W.; Slaughter, C. D.; Pierce, A. K.; and Aikens, R. S.: True Intensities of Fraunhofer Lines. Solar Phys. 18, 1971, pp. 366-378.

5. Brault, J. W. and White, O. R.: The Analysis and Restoration of Astronomical Data via the Fast Fourier Transform. Astron. and Astrophys. 13, 1971, pp. 169-189.

6. White, O. R. and Livingston, W.: Solar Luminosity Variation. II. Behavior of Calcium H and K at Solar Minimum and the Onset of Cycle 21. Ap. J. 226, 1978, pp. 679-686.

7. Chapman, G. A. and Sheeley, N. R., Jr.: The Photospheric Network. Solar Phys. 5, 1968, pp. 442-461.

8. Livingston, W. and Holweger, H.: Solar Luminosity Variation. IV. The Photospheric Lines, 1976-1980, submitted to Ap. J.

9. Mazzitelli, I.: Solar Models, Helium Content and Mixing Length. Astron. and Astrophys. 79, 1979, pp. 251-253.

10. White, O. R. and Livingston, W. C.: Solar Luminosity Variation. III. Calcium K Variation from Solar Minimum to Maximum in Cycle 21, submitted to Ap. J.

TABLE I. PHOTOSPHERIC LINE PARAMETERS AND OBSERVED 5-YEAR CHANGE

	Wavelength (A)	Equivalent Width W (mA)	EP (eV)	Landé g	$\Delta W/W^*$ K^{-1} (%)	$\Delta W/W^{**}$ obs. (%)
Fe I	5249.1	42.	4.5	0.9	-0.07	-0.45
Fe I	5250.2	74.	0.1	3.0	-0.09	-0.50
Fe I	5250.6	108.	2.2	1.5	-0.05	-0.17
Fe I	5379.6	62.	3.7	1.0	-0.05	-0.60
C I	5380.3	22.	7.7	1.0	+0.12	-2.30
Ti II	5381.0	65.	1.6	0.9	-0.00	-0.71
Si I	10827.1	423.	4.9	1.5	-0.04	0.00

* Predicted from HSRA model

** Total change over 5 years

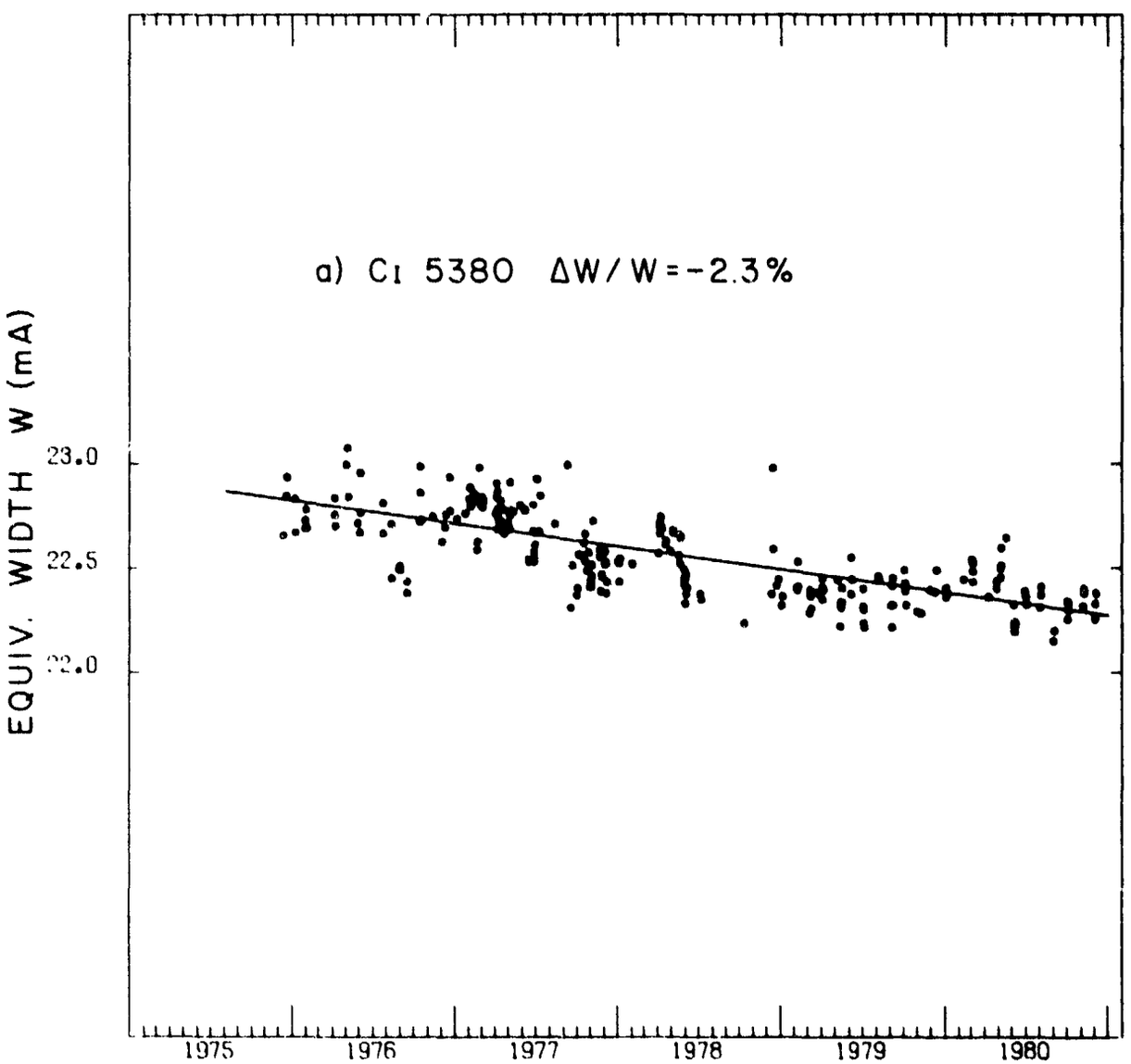


Figure 1a. Observed equivalent width in mA over 4-5 year time span: CI 5380

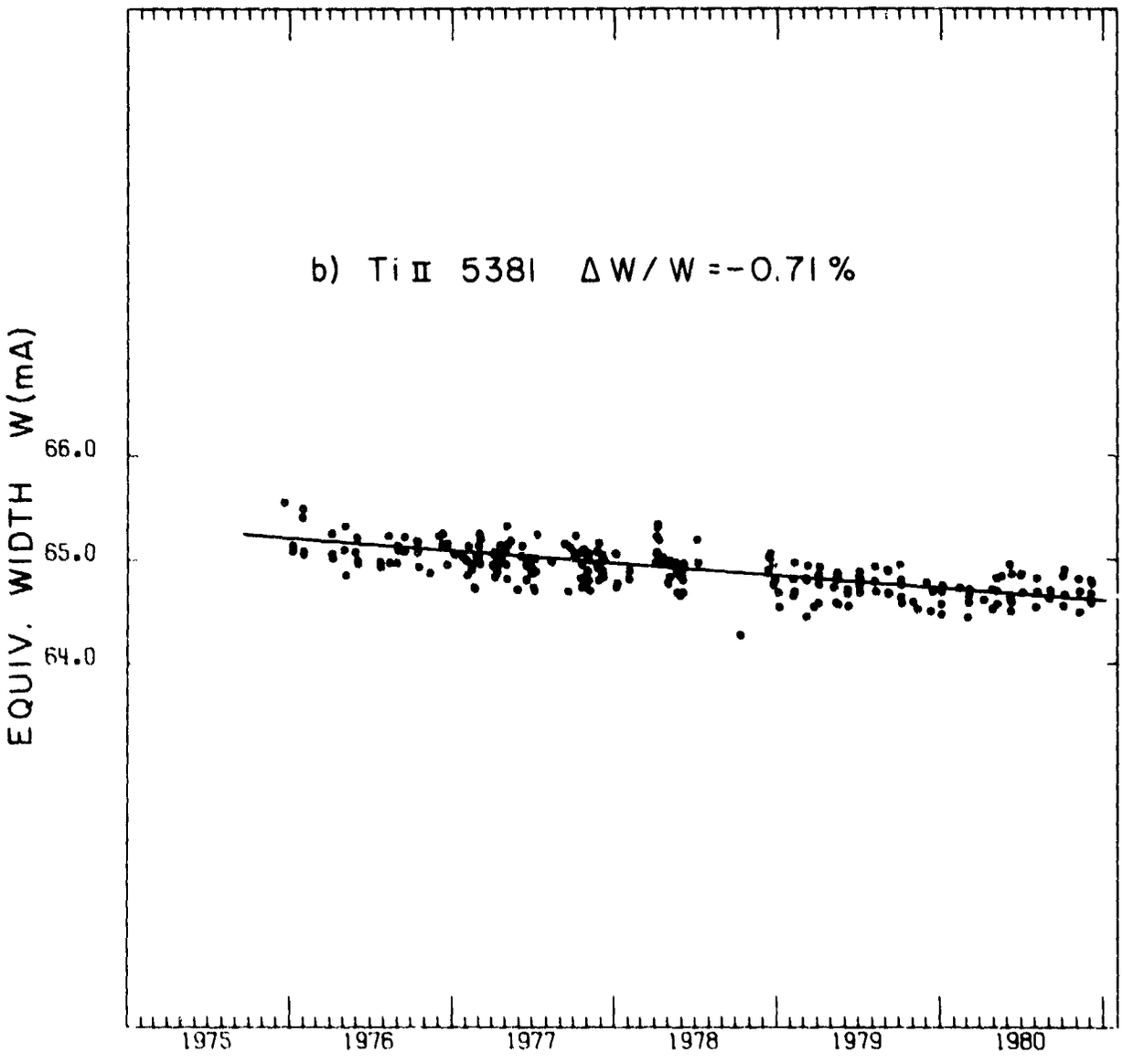


Figure 1b. Observed equivalent width in mA over 4-5 year time span: Ti II 5381

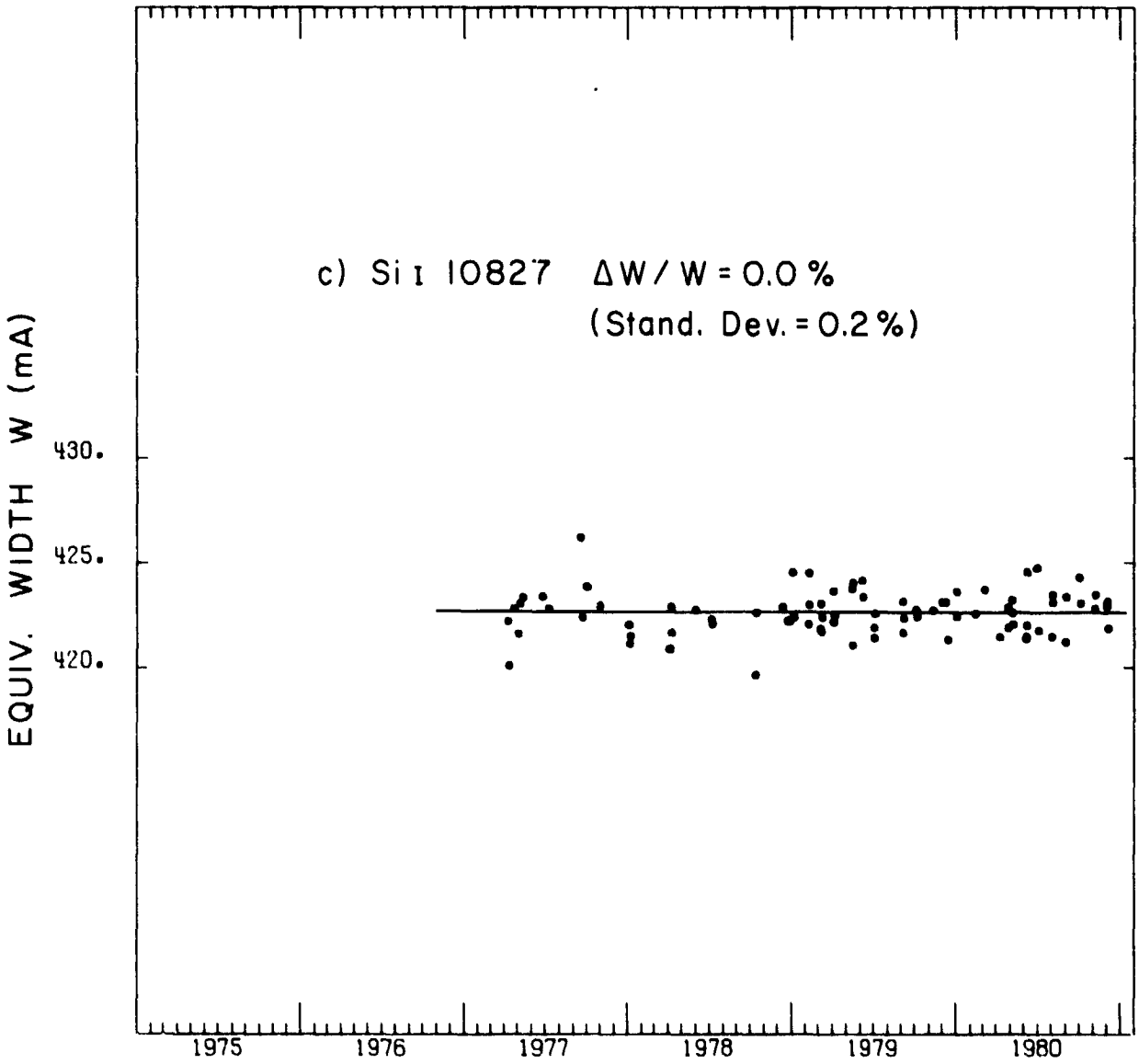


Figure 1c. Observed equivalent width in mA over 4-5 year time span: Si I 10827

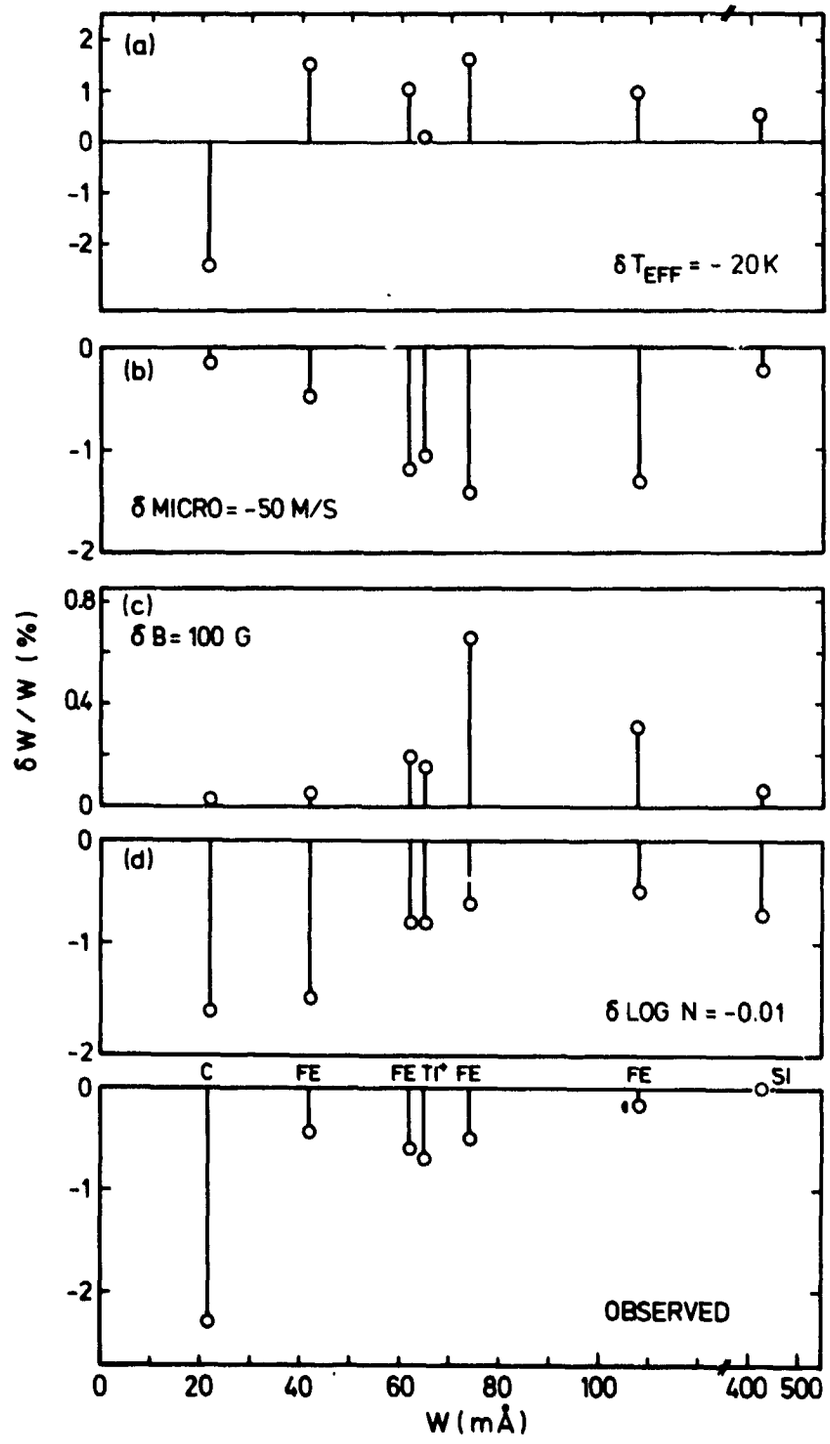


Figure 2. Predicted ΔW/W (%) compared with observed ΔW/W (%) for differing mechanisms: a) a step change of temperature, b) a change of microturbulence, c) introduction of a 100-gauss uniform magnetic field, and d) a 2.3% reduction in the abundance of observed species, and (bottom) the observations.

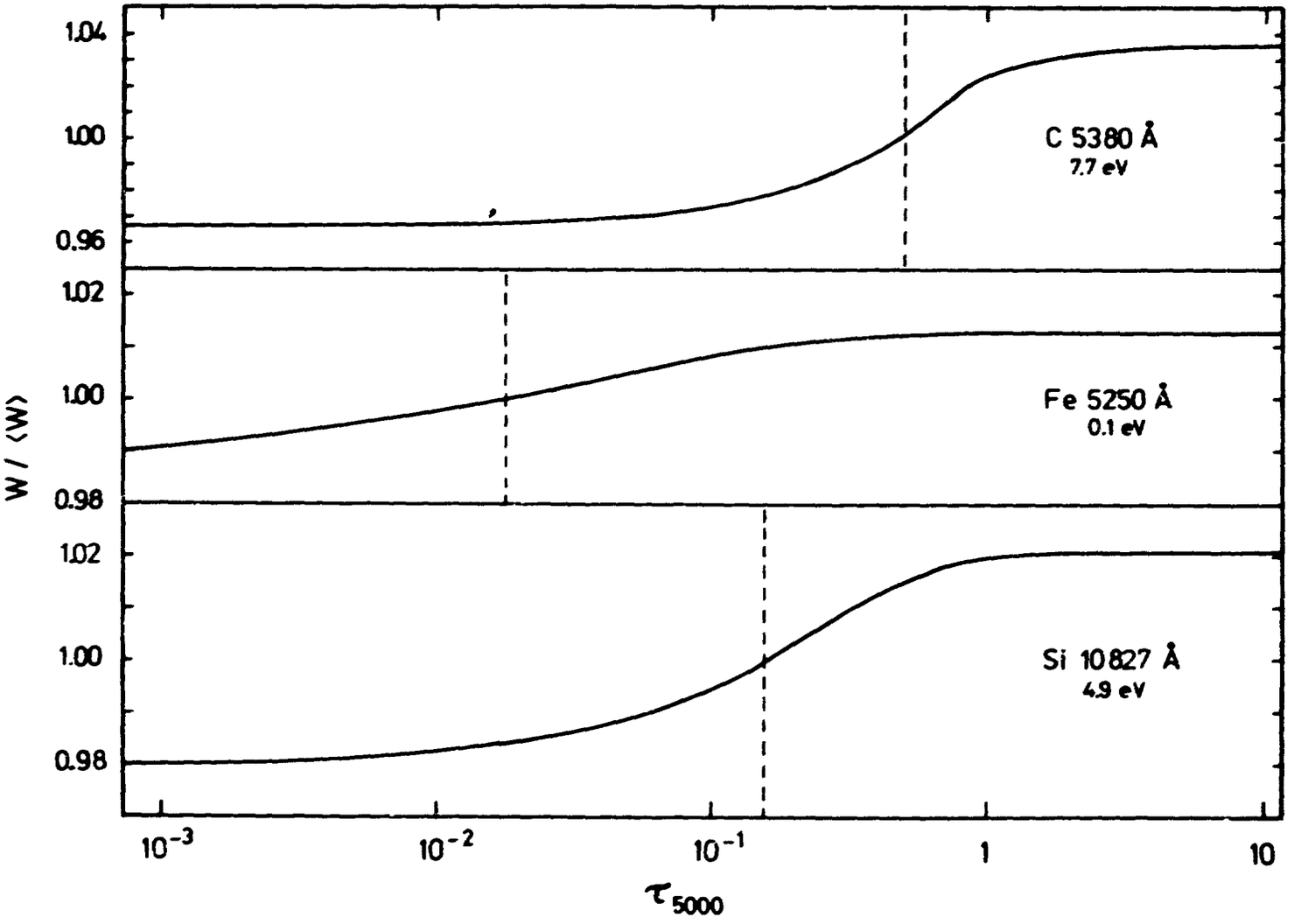


Figure 3. Contribution functions for the carbon, iron and silicon lines

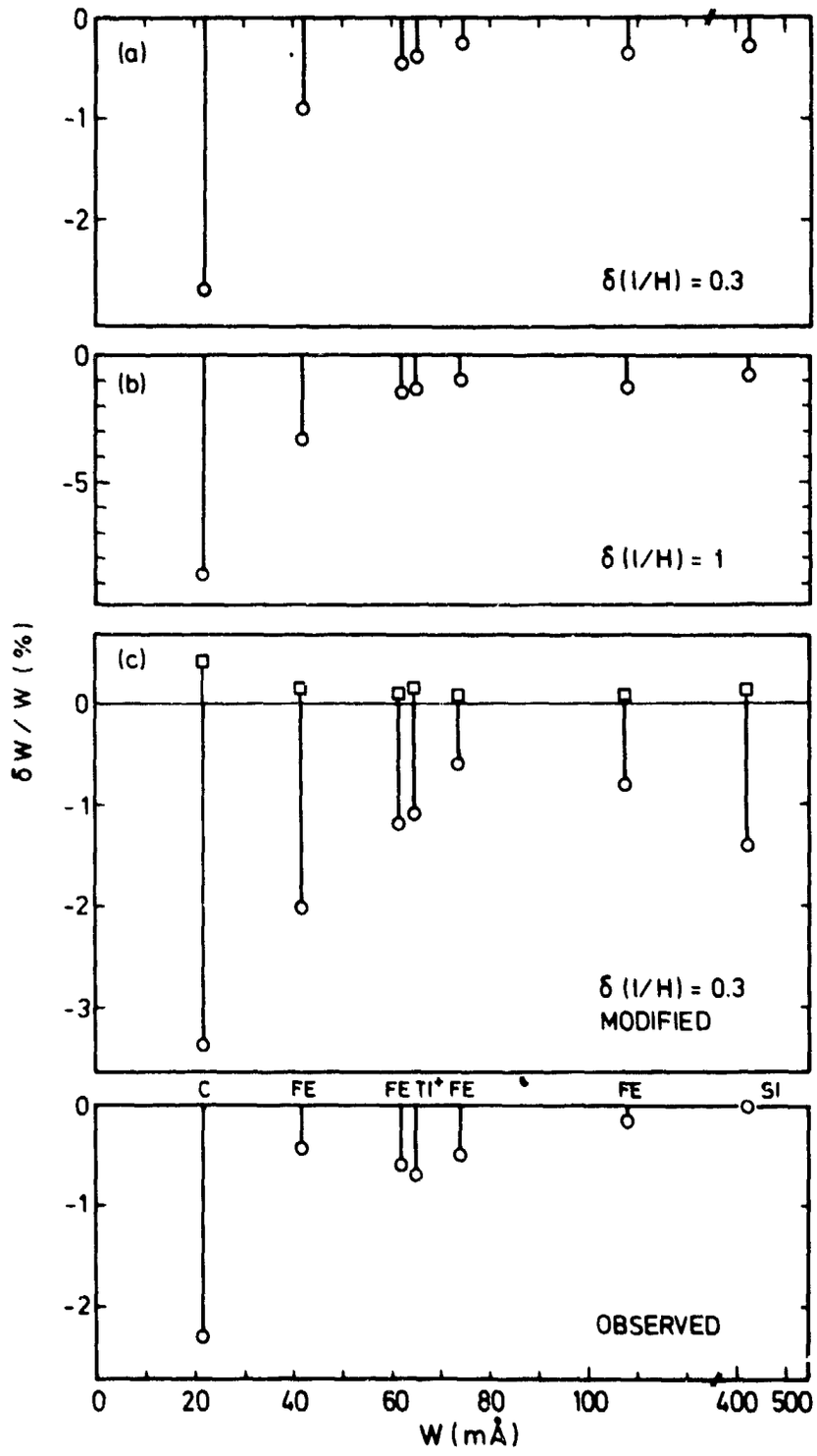


Figure 4. Equivalent width change in percent due to an increase of l/H a) from 2.0 to 2.3, b) 2 to 3, c) same as (a) but corresponding $\Delta T(\tau_{5000})$ shifted on τ scale by factor of two towards larger (squares), or smaller (circles) depths

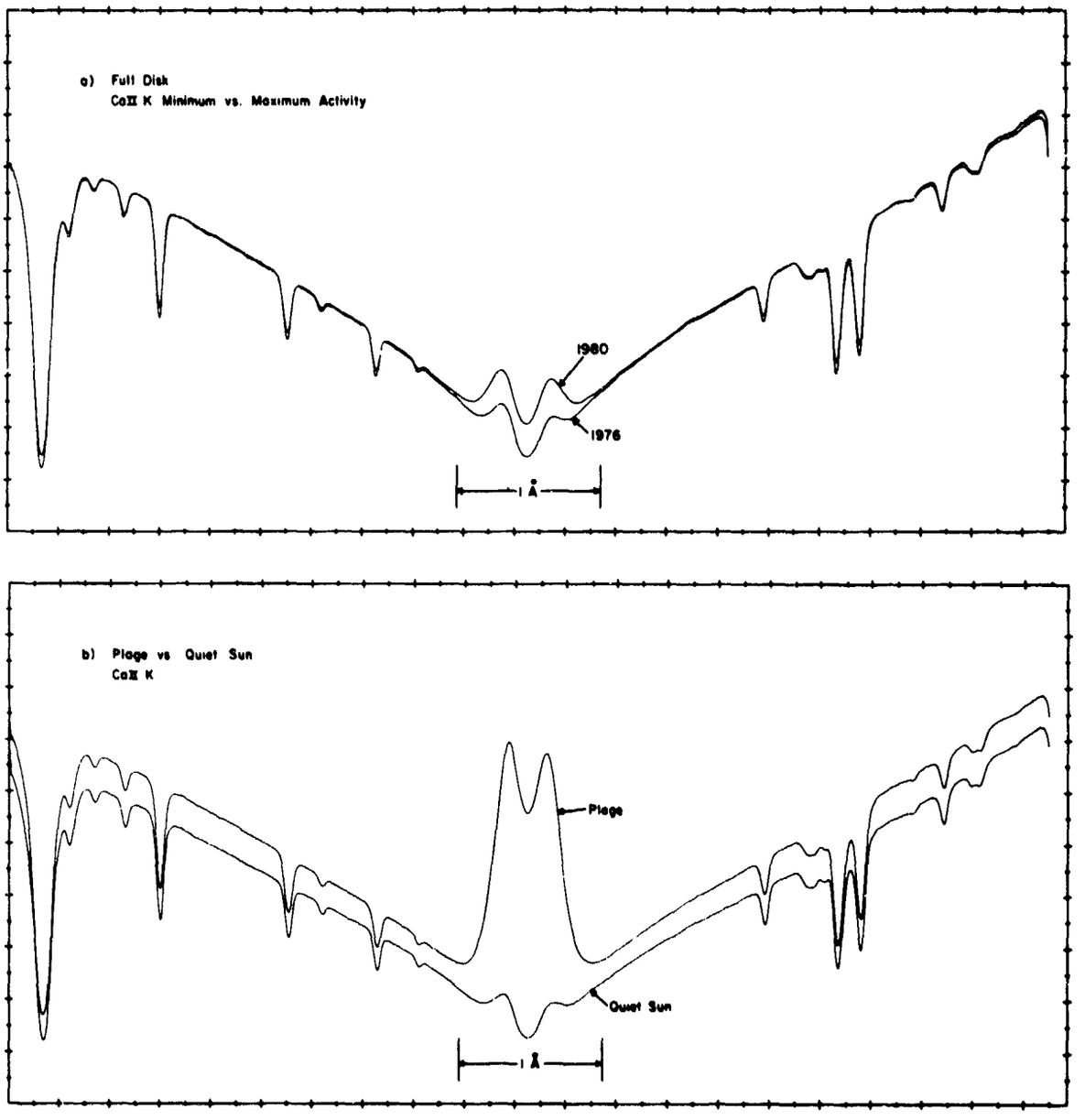


Figure 5. The variability of calcium K profiles: a) full disk K at minimum and maximum superposed and b) average profiles for an active region (plage) and a nearby quiet region at the same limb distance.

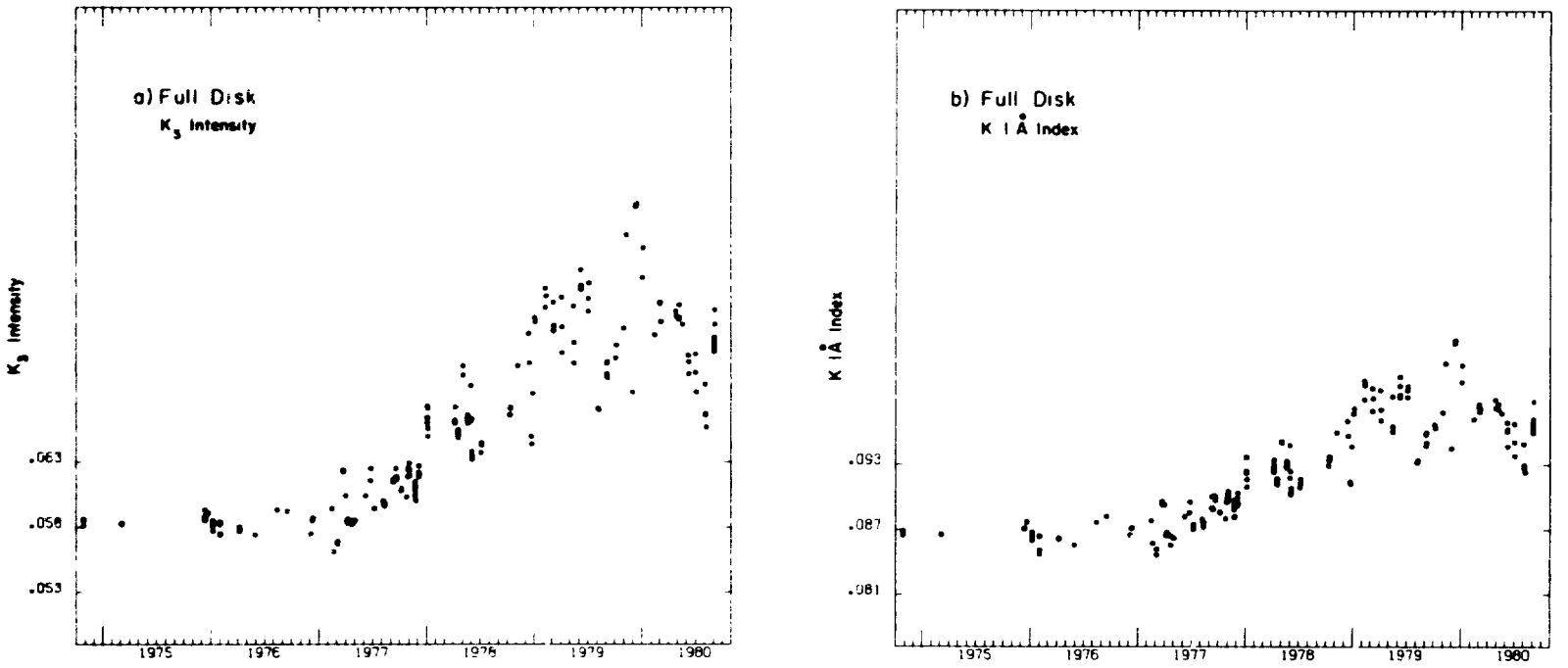


Figure 6. The cycle variation for K parameters: a) full disk K_3-intensity and b) full disk 1A K-index

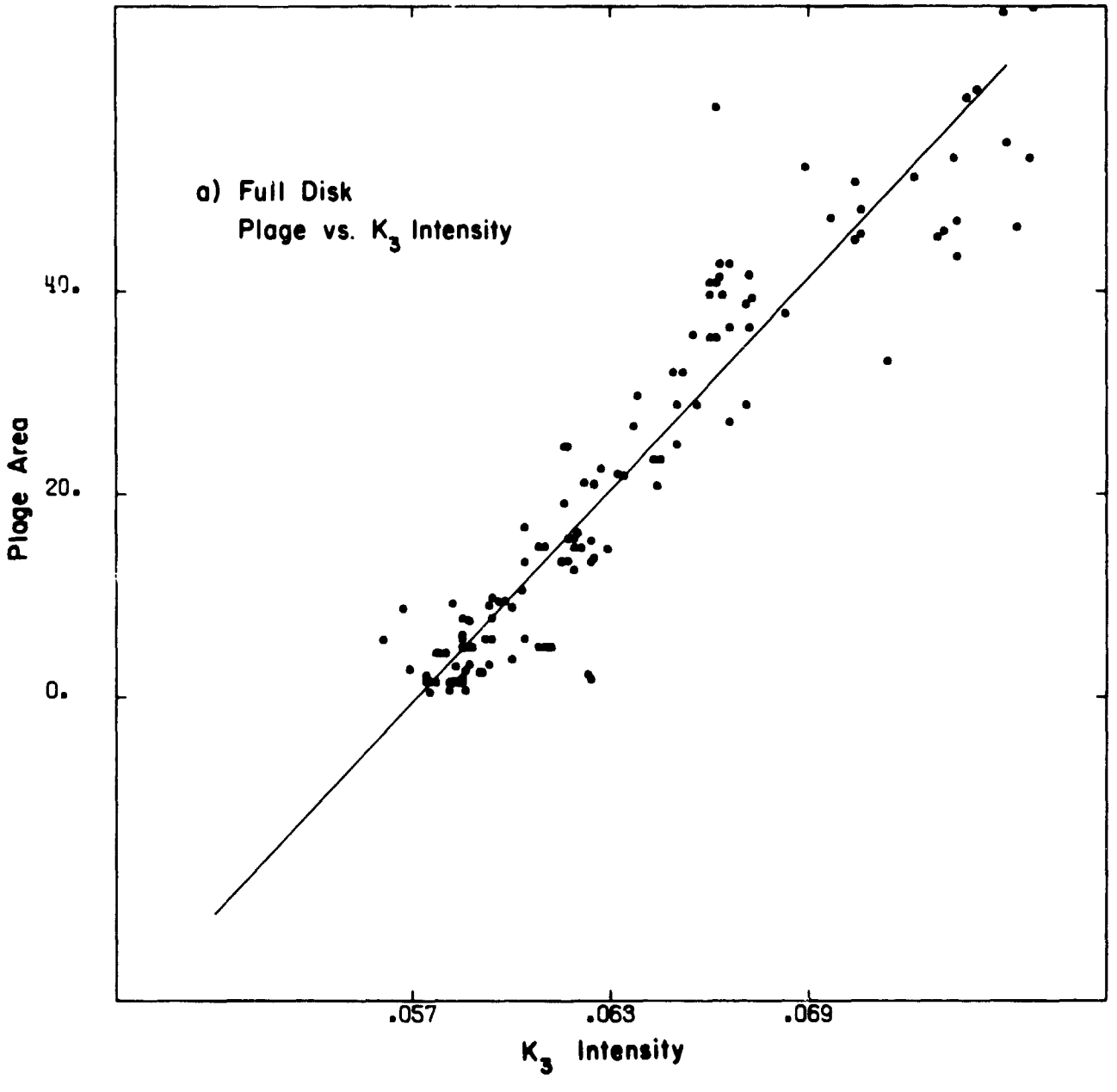


Figure 7. Correlation between K_3-intensity and the Mt. Wilson plage area index

THE HORIZONTAL AND VERTICAL SEMI-DIAMETERS OF THE SUN OBSERVED AT THE CAPE OF GOOD HOPE (1834-1887) AND PARIS (1837-1906): A REPORT ON WORK IN PROGRESS*

Clayton Smith and Dan Messina†
U.S. Naval Observatory

ABSTRACT

Cape and Paris meridian observations of the solar limbs which permit an estimate to be made of the solar semi-diameter are being surveyed, sampled, and compared with Greenwich and U.S. Naval Observatory observations. Significant systematic errors have been found in the Paris work and have been correlated with changes of instruments and observers. It is unlikely that further work on the Paris series would shed light on the problem of changes in the solar semi-diameter. Preliminary results from the more stable Cape series indicate that work should continue on the compilation of data from Cape observations of the sun.

INTRODUCTION

The possibility of a secular decrease in the apparent solar semi-diameter (referred to standard conditions) has been suggested from studies of meridian circle observations made at the Royal Greenwich and the U.S. Naval Observatories (refs. 1,2,3). Two other series of observations not previously discussed are available from the Royal Observatory at the Cape of Good Hope (Cape) in South Africa from 1834 to 1887 (ref. 4) and at the Paris Observatory (Paris) from 1837 to 1906 (ref. 5). The Cape series is of particular interest because of the Cape's southern hemisphere location. The Paris series is of interest because of the 70 year time-span.

A method for reducing raw observations to standard conditions is given, the method was applied to selected subsets of the original mass of observational material, and the results are discussed.

DATA ANALYSIS

Solar observations were reduced as follows (ref. 6):

$$SD_H = \frac{15R\Delta\alpha\cos\delta}{2S}, \qquad (1)$$

where SD_H = horizontal semi-diameter at unit distance (one A.U.),
R = earth-sun distance in units of A.U. at the time of observation,

* NASA Contract PCN 961-72273(1C)
† University of Maryland

$\Delta\alpha$ = measured difference in time between the east and west limbs,
S = a correction factor for the sun's motion in right ascension during the time between meridian passage of the east and west limbs. $S = 1/(1-\Delta a^s/3609.86)$, where Δa^s is the rate of change in right ascension of the sun in units of seconds of time per mean solar hour. (See table I for the monthly value used in any year.), and
δ = sun's apparent observed declination.

In some cases the north polar distance (NPD) was given, rather than the declination. In those cases, sin(NPD) was substituted for cosδ.

In the other coordinate:

$$SD_V = \frac{R\Delta\delta}{2}, \qquad (2)$$

where SD_V = vertical semi-diameter at unit distance (one A.U.), and

$\Delta\delta$ = measured difference between north and south limb declinations corrected for refraction.

Our strategy was to survey several years at the beginning and end of an instrumental series. Annual averages of SD_H and SD_V have been computed for Cape for the years 1834, 1884-1887, and 1861-1865 and are summarized in table II. Annual averages of SD_H and SD_V from Paris for the years 1837-1841, 1859-1867, and 1885-1890 are summarized in table III.

The entire Cape series was observed with no significant change in instrumentation or observers. However, the Paris series is composed of subsets of observations with four diffent instruments as indicated in table III. Significant changes in the observing staff from one year to the next were also noted.

PRELIMINARY RESULTS

CAPE

A test of the Cape results SD_H and SD_V of table II for a linear rate with time, T, by means of a least squares fit yielded the solutions $d(SD_H)/dT = -0.6 \pm 0.6$ seconds of arc per century and $d(SD_V)/dT = -0.4 \pm 0.4$ seconds of arc per century from 1834 to 1887, and 1861 to 1887, resp. From a statistical point of view, these results can barely be regarded as significant. However, since the mean errors are of the same order of magnitude as the rates and not very much larger, and since the two independent solutions are in better agreement with each other than expected from their relative errors, there is some indication that a clearer picture may emerge if the survey of Cape sun observations is broadened to include data from the time interval 1865 to 1884, and close attention is paid to the change in the mix of observers from one year to the next.

This preliminary result may be compared with the results of (1) Eddy and Boornazian (refs. 2,3) who found a secular decrease of -2" per century in

SD_H and -0".8 per century in SD_V from Greenwich and U.S. Naval Observatory meridian observations;(2) I. Shapiro (ref. 7) who from transits of Mercury found that any decrease in the solar diameter is likely to be under 0".3 per century;(3) D. Dunham, et al. (ref. 8) who found from an analysis of solar eclipses that the solar radius has contracted 0.34 ± 0.2 seconds of arc in 264 years; and(4) A. Wittman (ref. 9) who from the agreement between the mean of Tobias Mayer's observations of the sun, 1756-1760, and recent photoelectric results obtained in the 1970's finds no support for a secular decrease in the solar radius.

PARIS

The Paris results are inconclusive. Large systematic differences of the personal equation of individual observers having an effect as large as two seconds of arc on the determination of SD_H have been documented, and explain the discordant values of SD_H for 1866,1867, and 1902-1906. On the other hand, the significant decrease of SD_H from the 1840's to the 1860's is consonant within their relative errors with a similar decrease in the Greenwich results.

There is no significant change in the Paris SD_V of the 1837-1841 period compared with the 1859-1863 period, which is not in agreement with the Greenwich results over the same interval of time. Since different instruments were used at Paris in the 1837-1841 and 1859-1863 periods, i.e., the Fortin Mural Circle was used in the first period and the Gambey Mural Circle was used in the second period, systematic instrumental effects probably are at the root of the disagreement between Paris and Greenwich over that interval of time.

It was very disturbing to find that for the subset of observations made with the Grande Instrument Meridienne from 1863 to 1906 for which we have values of SD_V, the values were systematically larger than the earlier Paris values by about 1.5 seconds of arc, and also systematically larger than the Greenwich SD_V by about the same amount in the interval 1863 to 1906. This abrupt change in the system was probably caused by an instrumental change rather than an observer change. We have been able to document that changes of observer from one year to the next which grossly affect the SD_H (compare SD_H values 1885-1889 to SD_H values 1902-1906) cause no significant change in the corresponding SD_V values observed with the same instrument.

FUTURE WORK

Our next efforts will focus on completing the discussion of the Cape observations and then turning to the long series of the U.S. Naval Observatory. We hope to use concurrent Naval Observatory observations of the limbs of Jupiter and Saturn to indicate how diameter measurements can be affected systematically by personal equation and changes in instrumentation apart from changes which may occur as the result of severe punishment of the instrumentation during solar observations.

REFERENCES

1. Gething, P.J.D.: Greenwich Observations of the Horizontal and Vertical Diameters of the Sun. Mon. Not. Roy. Ast. Soc., vol. 115, no. 5, 1955, pp. 558- 570.

2. Eddy, J.A. and Boornazian, A.A.: Secular Decrease in the Solar Diameter. Bull. of the Amer. Ast. Soc., vol. 11, no. 2, 1979, p. 437.

3. Lubkin, G.B.: Analyses of Historical Data Suggest Sun is Shrinking. Physics Today, vol. 32, no. 9, Sept. 1979, pp. 17-19.

4. Royal Observatory, Cape of Good Hope, South Africa. Annual volumes of meridian observations 1834-1877.

5. Paris Observatory Annals. Annual volumes 1837-1906.

6. Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac., London. Her Majesty's Stationery Office., 1961, pp. 137-141.

7. Shapiro, I.: Is the Sun Shrinking? Science, vol. 208, no. 4, April 1980, pp. 51-53.

8. Dunham, D.W.; Sofia, S.; Fiala, A.; Herald, D.; Muller, P.M.: Observations of a Probable Change in the Solar Radius Between 1715 and 1979. Science, vol. 210, no. 12, Dec. 1980, pp. 1243-1245.

9. Wittman, A.: Tobias Mayer's Observations of the Sun: Evidence Against A Secular Decrease of the Solar Diameter. Solar Physics, vol. 66, 1980, pp. 223-231.

TABLE I. S, A CORRECTION FACTOR FOR THE MOTION OF THE SUN

Jan.	1.00297	July	1.00281
Feb.	1.00271	Aug.	1.00259
Mar.	1.00253	Sept.	1.00249
Apr.	1.00256	Oct.	1.00258
May	1.00274	Nov.	1.00284
June	1.00288	Dec.	1.00307

TABLE II. ANNUAL MEANS OF SOLAR SEMI-DIAMETERS FROM OBSERVATIONS AT CAPE

Year	SD_H	No. of Obsns.	SD_V	No. of Obsns.
1834	961".50	132	--	0
1861	961.25	61	962".35	37
1862	960.63	37	962.04	32
1863	961.19	42	962.43	33
1864	961.57	54	962.29	53
1865	961.50	68	962.13	65
1884	960.86	69	962.09	77
1885	961.16	19	962.17	19
1886	961.34	141	962.27	150
1887	961.09	175	962.09	179

TABLE III. ANNUAL MEANS OF SOLAR SEMI-DIAMETERS FROM OBSERVATIONS AT PARIS

Year	SD_H	No. of Obsns.	SD_V	No. of Obsns.
Lunette Meridienne de Gambey(LMG)			Fortin Mural Circle	
1837	962".24	146	960".56	20
1838	962.62	111	960.70	4
1839	962.26	121	--	0
1840	962.75	142	961.35	11
1841	962.03	93	960.92	17
	LMG (continued)		Gambey Mural Circle	
1859	961.90	136	--	0
1860	961.20	77	961.28	3
1861	961.55	108	960.49	4
1862	960.67	42	961.56	8
1863	960.57	58	961.25	36
Grande Instrument Meridienne (in both coordinates)				
1863	961.19	25	963.18	20
1864	961.54	102	962.45	141
1865	960.62	101	961.71	98
1866	962.68	54	962.67	45
1867	962.03	76	962.50	73
1885	961.20	114	962.39	96
1886	961.25	126	962.44	119
1887	961.68	92	962.56	90
1888	961.29	93	962.45	57
1889	961.48	173	962.54	79
1902	962.85	105	962.42	93
1903	962.87	93	962.29	92
1904	962.80	97	962.49	89
1905	962.92	88	962.52	82
1906	963.05	88	962.54	87

D11

ECLIPSE RADIUS MEASUREMENTS*

David W. Dunham and Joan Bixby Dunham
Computer Sciences Corporation

Alan D. Fiala
U. S. Naval Observatory

Sabatino Sofia
Goddard Space Flight Center

ABSTRACT

We have improved methods for predicting the path edges and reducing observations of total solar eclipses for determining variations of the solar radius. Recently-analyzed observations of the 1925 January eclipse show a 0".7 (arc second) decrease in the solar radius during the past fifty years.

INTRODUCTION

Methods for predicting the location of the edges of paths of total solar eclipses, methods of reducing timed observations made just inside the edges of the path, and accuracies of the solar radius variations determined from these data, are described elsewhere (ref. 1). Results for the eclipses of 1715, 1976, and 1979 have been reported (ref. 2).

IMPROVED RESULTS

Our previous procedure for computing the non-spherical corrections for eclipse path edges (ref. 1) has been replaced by a non-iterative scheme and has been largely automated. Similarly, we have written computer programs to efficiently reduce reported Baily bead and contact timings, permitting more comprehensive analyses of the data. The method of reduction has been described elsewhere (ref. 3). For a reported timing, the computer programs print a plot showing the lunar limb derived from a U. S. Naval Observatory magnetic disk file of limb correction data (ref. 4), the solar limb, and a pointer indicating the lunar feature which probably caused the event.

*This work was supported by N.A.S.A. Grant NAS-5-24350, N.S.F. Grant ATM-77-23757, and the U. S. Naval Observatory.

The selected lunar feature can be replaced with another choice, if this seems appropriate after examining the plots. See Fig. 1.

Besides the above improvements, some of our other recently-published results have been superceded because the observed events used in the analysis were not restricted to position angles within 30° of the lunar axis of rotation (ref. 5 and 6), in the polar regions where our arguments about minimal libration-dependent errors are valid, or because some of the computed results were misinterpreted (ref. 7).

THE 1925 ECLIPSE

As a result of E. W. Brown's efforts, the total solar eclipse of 1925 January 24 was well-observed from sites near both the northern and southern limits (ref. 8). Brown noted that the observations determined the location of the actual path edges very accurately, but that the result could not be used due to lack of knowledge of the location of the lunar valley bottoms, which caused the observed contacts, with respect to the moon's center of mass. Although the southern-limit observations made from New York City were published (ref. 9), the reports from the northern limit collected by Brown have apparently been lost. Fortunately, we found a report made at a professional observatory situated very close to the northern limit (ref. 10). We have measured the positions of the observers from 1:24,000-scale topographic maps of the U. S. Geological Survey. These positions are accurate to about 1" in geographic latitude, which contributes an error of only 0".01 to our determination of the solar radius from the observations.

The correction to the standard solar radius (959".63 at a distance of 1 astronomical unit) derived from the 1925 eclipse, along with improved values for the corrections for previously-studied eclipses, and the standard errors of the corrections, are given in Table 1. The lunar valleys which produced the observed contacts at the limits of the 1925 eclipse also produced events which were timed in 1979 (two contacts and one Baily bead event) and in 1980 (three Baily bead events). The solar radius corrections derived from solutions using only these observations are also listed.

CONCLUDING REMARKS

While no significant change in the solar radius can be deduced from observations of recent eclipses, the data show that the radius decreased by at least 0".5 (0.05% or 350 km) between 1925 and 1979. These two eclipses have special significance since they are three Saros cycles apart, producing similar geometries, especially similar lunar librations. The decrease is 0".70 ±0".13 if all polar timings are used. This result is confirmed (0".51 ±0".24), to less precision, if only those lunar features producing observed events during both eclipses are used. Unlike the result for the 1715 eclipse, there is no significant uncertainty in the coordinates of the observers of this century's eclipses. The physical significance of variations of the solar radius is discussed in other papers in these proceedings.

REFERENCES

1. Sofia, S.; Dunham, D. W.; and Fiala, A. D.: Determination of Variations of the Solar Radius from Solar Eclipse Observations. Proc. Conf. Ancient Sun, 1980, pp. 147-157.

2. Dunham, D. W.; Sofia, S.; Fiala, A. D.; Herald, D.; and Muller, P. M.: Observations of a Probable Change in the Solar Radius Between 1715 and 1979. Science, vol. 210, no. 4475, 12 Dec. 1980, pp. 1243-1245.

3. Herald, D.: Observations of Baily's Beads from Near the Northern Limit of the Total Solar Eclipse of June 20, 1974. The Moon, vol. 16, 1976, pp. 91-100.

4. Watts, C.B.: The Marginal Zone of the Moon. Astron. Papers Amer. Ephem., vol. 17, 1963.

5. Fiala, A. D.; Dunham, D. W.; and Dunham, J. B.: Preliminary Results from Observations made near the Edges of the Path of the 1980 February 16 Total Solar Eclipse. Bull. Amer. Astron. Soc., vol. 12, no. 2, 1980, p. 474.

6. Fiala, A. D.; Dunham, D. W.; and Dunham, J. B.: Observations of Baily's Beads from the Edges of the Path of the Total Solar Eclipse of 1980 February 16. Bull. Indian Astron. Soc., vol. 8, no. 2&3, 1980, pp. 81-82.

7. Dunham, J. B.; Dunham, D. W.; and Fiala, A. D.: Solar Radius Variations from Observations of Four Eclipses. Bull. Amer. Astron. Soc., vol. 12, no. 4, 1980, p. 832.

8. Brown, E. W.: Discussion of Observations of the Moon at or near the Eclipse of 1925 January 24. Astron. J., vol. 37, 1926, pp. 9-19.

9. Magalhaes, F. V.; Law, C. L.; and Lieb, J. W.: Observations of the Total Solar Eclipse of January 24, 1925, made by Electric Companies Affiliated with the Consolidated Gas Company of New York. Trans. Illumin. Eng. Soc., vol. 20, no. 6, 1925, pp. 565-628.

10. Seagrave, F. E.: Short Eclipse Notes: At North Scituate, Rhode Island. Pop. Astron., vol. 33, 1925, p. 279.

TABLE I. SOLAR RADIUS CORRECTIONS DETERMINED FROM OBSERVATIONS NEAR ECLIPSE PATH EDGES

Date	Timings within 30° of lunar axis		Timings using features determining 1925 events		Lunar librations	
	No.	$\Delta r_\odot$	No.	$\Delta r_\odot$	long.	lat.
1715 May 3	4	+0".52 ±0".2	0	--	+1°.8	-0°.2
1925 Jan. 24	4	+0.62 ±0.08	4	+0".62 ±0".08	+2.5	-0.2
1976 Oct. 23	15	-0.23 ±0.14	0	--	-1.4	+0.1
1979 Feb. 26	33	-0.08 ±0.09	3	+0.11 ±0.23	+1.7	-0.3
1980 Feb. 16	55	-0.03 ±0.04	3	-0.05 ±0.35	-3.0	-0.1

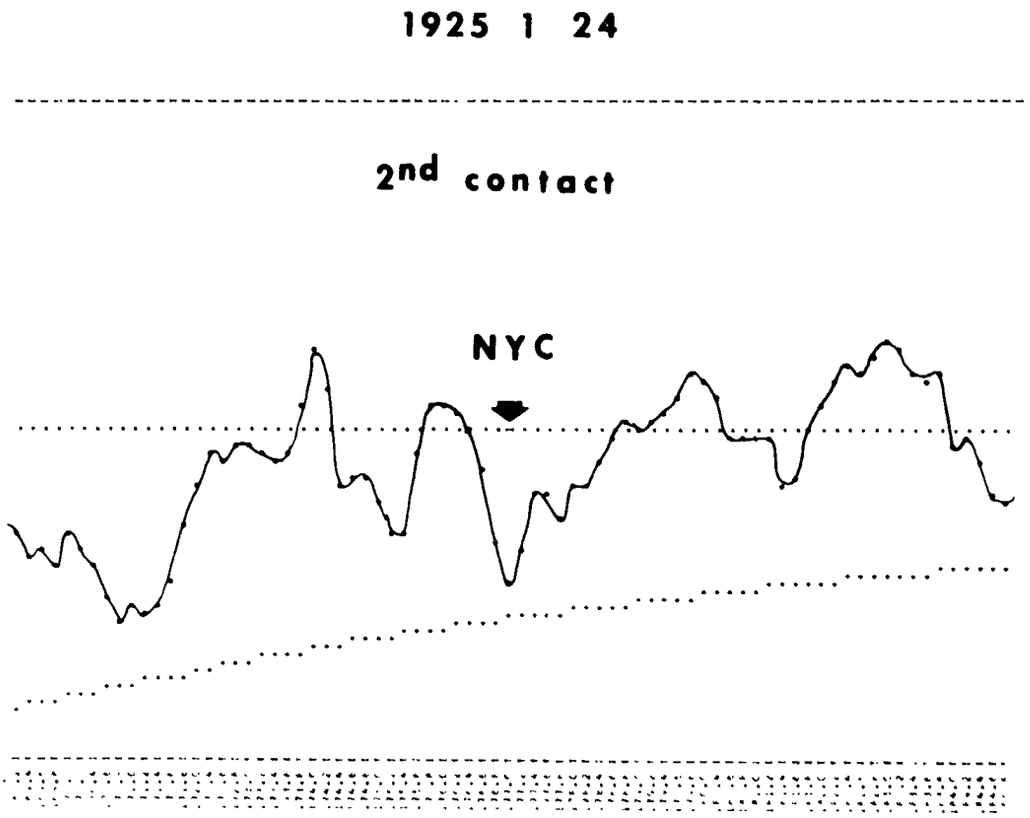


Figure 1. This printer plot shows the predicted positions of the lunar and solar limbs computed for the time of second contact for the observer in New York City closest to the southern limit of the 1925 January 24th total eclipse path. The horizontal line of dots marks the Moon's mean limb. Sideways I's at the top and bottom indicate distances 4" above and below the mean lunar limb, respectively. Asterisks connected with a hand-drawn curve show the actual lunar profile according to Watts' charts (ref. 4). Pluses mark the solar limb. The vertical scale is exaggerated 27 times relative to the horizontal scale. Position angles measured in degrees from the Moon's northern axis as defined by Watts are given at the bottom. The dot in the mean limb line which is replaced by a sideways number 1 indicates the computer-selected choice of the valley which produced second contact, where the predicted separation of the actual lunar limb and the solar limb was minimized.

SCLERA SOLAR DIAMETER OBSERVATIONS[1]

Henry A. Hill, Thomas P. Caudell and Randall J. Bos
Department of Physics
University of Arizona
Tucson, Arizona 85721

ABSTRACT

The accurate measurement of the solar shape and diameter has proven to be a very difficult endeavor, as evidenced by the large scatter in the data record. Although observer bias, changing environment and atmospheric seeing have been identified as the major sources of this scatter, attempts to correct for these still leave disagreement regarding the constancy of the solar diameter. Most data have been derived from solar meridian and Mercury transit timings, with total solar eclipse observations adding some additional information. Interest in the time dependence of the mean solar diameter has been revived lately as a means to infer indirectly variability in the solar luminosity. However, theoretical results regarding the relationship between these two properties are strongly model-dependent, demonstrating the need for independent measures of diameter and luminosity variations before inferences about the latter can be made. The recent use of computer-controlled photoelectric devices has led to the systematic removal of observer bias and atmospheric seeing from measurements of the solar diameter, but a new set of problems has surfaced due to variations in the sun itself. Careful attention must be given to the definition of an edge on the solar disk in order to distinguish between observed diameter variations due to physical shape changes and those due to variability in other solar properties which affect the edge definition. For example, the discovery at SCLERA[2] of differences in the solar limb darkening function between equator and pole required that special precautions be taken to extract a visual oblateness determination with minimal interference from such effects. Therefore, programs which propose to accurately measure intrinsic solar diameters must not only guard against the classical systematics mentioned earlier, but must be prepared to detect changes in other properties of the sun at the extreme limb.

---

[1]The work was supported in part by the National Aeronautics and Space Administration and the Department of Energy.

[2]SCLERA is an acronym for the Santa Catalina Laboratory for Experimental Relativity by Astrometry, a facility jointly operated by Wesleyan University and the University of Arizona.

INTRODUCTION

With regard to man's apparent desire for stability, i.e., constancy, in the structure of the universe, the sun has been a considerable disappointment. Variability seems to occur in each observed solar property when it is examined with sufficient care. The advent of the telescope brought about the discovery of the sunspot cycle; later the magnetic cycle was found and recently, with the addition of new observing techniques, variation has been detected in the apparent solar shape. Focus is now being given to possible variations in solar luminosity and accurate methods of monitoring it. Aside from direct bolometry, one methodology for this type of research makes use of measurements of the solar diameter and limb darkening function as indirect indicators of the solar luminosity. This approach, which is currently being used at SCLERA, will be reviewed here.

DIAMETER VS. LUMINOSITY VARIATION

A change in the solar diameter will occur when a change in energy output or a redistribution of energy within the sun occurs. Some recent theoretical work has been directed toward attempting to identify the fundamental physical processes affecting the relationship between a change in radius and a change in luminosity. Reference 1 and, in more detail, reference 2 describe work to derive such a relationship by varying the convective efficiency used in local mixing length theory. Reference 3 uses changes in magnetic flux tube buoyancy to alter the radius while holding the luminosity constant. The argument is put forth in reference 4 that variations in lower convection zone magnetic fields with the solar cycle could alter the difference between the actual and adiabatic temperature gradients, thus affecting the convective flux. The results of these models vary considerably in the predicted magnitude of the relationship between radius and luminosity variation, even predicting different signs in some cases. Comparisons between the various models can be found in references 5 and 6. It appears at this time that independent observations of solar diameter and luminosity will be necessary to resolve these differences.

Another problem, and one which is commonly overlooked, relates to the observational definition of a solar diameter and its relationship to the theoretically modeled perturbations. A discussion of the use of the solar diameter as an indirect measure of luminosity makes little sense without careful consideration of this matter. We will deal with this in a later section. Nonetheless, historical records of diameter have been analyzed for variation and these analyses are reviewed in the following section.

HISTORICAL RECORD OF THE SOLAR DIAMETER

Attempts have been made for over three centuries to measure the diameter of the sun. These observations, which consist of meridional transits, Mercury transits, and solar eclipse timings, have been reviewed recently in an effort

to determine the long-term character of the solar diameter. Reference 7, in an analysis of the Greenwich meridian circle data, initially reported a decrease in the solar diameter of nearly 2" per century; this figure has recently been revised downward to about 0".1 per century (reference 6). An analysis of the record of Mercury transit timings from the late 1700's to the present has resulted in the conclusion, at the 90% confidence level, that the solar diameter has decreased by less than 0".3 per century and, within the stated error, probably not at all (reference 8). References 9 and 10 employed a large sample of published solar semidiameter determinations taken between 1660 and 1978 and found no indication of a substantial secular variation, although a variation with solar cycle may exist. Finally, timings of eclipses made near the limits of totality in 1715 and 1925 show that the sun was larger than it is currently by 0".3 and 0".7, respectively (references 11 and 12).

It is clear from the above summary that the historical record provides no clear interpretation. Changing environmental conditions, observer bias, and errors in correcting systematics like irradiance, lunar limb shape, and seasonal changes in refraction partially contribute to the "scatter" in the observations. The salient point, however, is this: the weight of these studies indicates that the solar diameter is variable at the 0.1% level over time scales of 10-100 years. Modern observations and techniques will be necessary for more detailed conclusions.

MODERN OBSERVATIONS

The use of photoelectric devices to determine the shape and diameter of the sun allows greatly increased precision. The usual observer biases are immediately removed and the systematics can be controlled for. Even so, controversy has arisen from the intercomparison of modern measurements of the solar shape (references 13 and 14), primarily because of differing definitions of the solar edge used in the diameter determinations. In fact, this intercomparison led to the discovery of changes in the solar limb darkening function on time scales of several months (reference 13). These changes are manifested as an excess equatorial brightening and were detected through the use of an edge definition which distinguishes between physical changes in radius (i.e., a rigid translation of the limb) and changes in the shape of the intensity pattern used to define the edge. This definition, which also demonstrates a markedly reduced sens ivity to seeing effects, is referred to as the finite Fourier transform definition or FFTD (see reference 15 for a more complete description). When the FFTD was applied in an observing program to measure solar oblateness at the SCLERA telescope during the early 1970's, the oblateness was found to contain a time-varying component. The properties of the FFTD allowed the differentiation of those periods of time when the excess equatorial brightness was minimal, i.e., when the equatorial and polar limb darkening functions were most similar, resulting in a determination of the intrinsic visual solar oblateness.

These measurements have far-reaching implications. Changes in the solar limb darkening function must be accompanied by changes in, for example, the temperature and density profiles throughout the photosphere and chromosphere. Alteration in these properties will change the energy flux associated with

various energy transport mechanisms. The existence of a time-varying excess equatorial brightness implies, then, an anisotropy in the solar flux. The accuracy to be expected for direct bolometric measurements of the solar constant will be strongly affected by this anisotropy, since any observation of that type works in a very narrow solid angle.

Variability of the limb darkening function may well pose the ultimate limitation on the precision to which a solar diameter can be determined, and progress must be made in understanding its effects on measurements of the solar diameter. On the other hand, changes in the limb darkening function, rather than in the absolute solar diameter, may turn out to be the dominant source of information on the solar driving function for the earth's climate.

SCLERA'S OBSERVING PROGRAM

At SCLERA, an active program is planned for 1980-81 (reference 8) to study long-term solar variability and its effect on solar energy output. For the first time the full astrometric capabilities of the SCLERA facilities will be utilized. A review of the various aspects of this program follows.

To accurately measure changes in the shape and diameter of the sun, a Michelson interferometer with a He-Ne laser source is used. This device continuously monitors the separation between the detectors which gather the limb intensity profiles used in the FFTD definition of the edge. To study long-term solar variability, extreme stability of this interferometer and its supporting hardware must be maintained throughout the planned observing period. To meet this goal for the short term, approximately the first year, the observing instrument will perform in a "frozen" configuration. This implies that no modifications or improvements will be made to the hardware during this initial period. To allow the comparison of diameter observations taken from day to day, a white light interferometer has been implemented which coexists with the Michelson interferometer. The measurement of the white light fringe with respect to the laser fringe provides the necessary day-to-day fiducial for the short term.

The balance between freezing an astrometric instrument to insure consistency of results and the necessity to respond to new technological developments, thus staying at the frontier of the field, represents a major dilemma in astrometry. A calibration scheme is currently being developed at SCLERA to avoid this predicament. The new device uses grating optics and laser light to produce reference fiducials in the telescope's focal plane at a standard angle of separation. This calibration system will allow absolute diameter measurements to be made with high precision in a historically repeatable manner. In addition, its incorporation into the observing program will remove the necessity to maintain the detectors in a frozen state and permit technological upgrading when appropriate.

Another feature of SCLERA's observing program is the continuous record of the limb darkening function which is made simultaneously with the diameter measurements. The FFTD edge definition is then applied to these intensity profiles off-line to generate a sensitive "barometer" of changes in the limb

darkening. This will provide the necessary information to separate physical diameter changes from those due to changes in the shape of the limb, thus avoiding the possible misinterpretations mentioned earlier.

The 1980-81 observing program calls for daily operation of the telescope, weather and equipment permitting. The analysis of these new observations should provide, among other things, an improved value for solar oblateness and a new measure of the long-term variability of the sun. It is anticipated that this program at SCLERA should launch a new era in the study of the global properties of the sun.

REFERENCES

1. Sofia, S., O'Keefe, J., Lesh, J. R., and Endal, D. S. Science, vol. 204, 1979, p. 1306.

2. Dearborn, D. S. P. and Blake, J. B. Astrophysical Journal, vol. 237, 1980, p. 616.

3. Thomas, J. H. Nature, vol. 280, 1979, 663.

4. Spiegel, E. A. and Weiss, N. O. Nature, vol. 287, 1980, p. 616.

5. Gilliland, R. L. Astrophysical Journal, 1980 (submitted).

6. Gough, D. Nature, vol. 288, 1980, p. 639.

7. Eddy, J. A. and Boornazian, A. A. American Astronomical Society Bulletin, vol. 11, 1979, p. 437.

8. Shapiro, I. I. Science, vol. 208, 1980, p. 51.

9. Wittmann, A. Astronomy and Astrophysics, vol. 61, 1977, p. 225.

10. Wittmann, A. Personal communication, 1981.

11. Dunham, D. W., Dunham, J. B., Fiala, A. D., and Sofia, S. Solar Constant Workshop. NASA CP-2191, 1981.

12. Dunham, D. W., Sofia, S., Fiala, A. D., Herald, D., and Muller, D. M. Science, vol. 210, 1980, p. 1243.

13. Hill, H. A. and Stebbins, R. T. Astrophysical Journal, vol. 200, 1975, p. 471.

14. Dicke, R. H. and Goldenberg, H. M. Physics Review Letters, vol. 18, 1967, p. 313.

15. Hill, H. A., Stebbins, R. T., and Oleson, J. R. Astrophysical Journal, vol. 200, 1975, p. 484.

16. Hill, H. A. Proceedings of Workshop "Solar Instrumentation -- What's Next," Sacramento Peak National Observatory, Sunspot, New Mexico, 1980.

SOLAR RADIUS MEASUREMENTS

T. L. Duvall, Jr. and H. P. Jones

NASA/GSFC, Laboratory for Astronomy
and Solar Physics
Solar Physics, Southwest Solar Station
Tucson, Arizona

ABSTRACT

Preliminary results of solar radius measurements made during 1979-1980 are discussed. Variability in the radius measurements of 0.4π is found, of unknown origin.

INTRODUCTION

An observational program to search for variability of the solar radius was begun in December 1979 at Kitt Peak. This program was initiated because of the claims of a secular decrease in the solar radius of ~ 1π/century (ref. 1). A search of the literature, however, also reveals claims of variability on time scales such as a 0.1π solar cycle variation (ref. 2). The first ten months of observations are discussed in this preliminary report. Unexplained variability at the level of 0.4π is found. At this stage of the investigation, all possible sources for this variability are being considered—instrumental, atmospheric, and solar.

OBSERVATIONS

The technique used to measure the solar radius is the classical one of obtaining the time interval between passage of opposite solar limbs across a detector pointed at a fixed place in the sky. The known rotation rate of the earth and declination of the sun are then used to derive an angular diameter. In the present application of this technique a linear array of 512 detectors is used. Each detector subtends an angle of 1π x 1π. The detector array, aligned at a 15° angle with respect to the solar drift direction, is used to measure 512 chord lengths in the neighborhood of the diameter of the circular solar disk. A diameter is derived from the chord lengths by a least-squares fit. The telescope used is the Vacuum Solar Telescope of the Kitt Peak National Observatory. The detector array and associated electronics is part of the 512 channel magnetograph normally used with this telescope.

An observation consists of recording the signals from the detector array with a time constant of 0.3 s for the time it takes the sun to drift across the entire array (~ 2 1/2 minutes). Four such observations were made on each of 18 days between 4 December 1979 and 30 September 1980. For each day's observations radius values are computed. The mean value of each set of 4 and also a statistical uncertainty are derived. The statistical uncertainties are consistent with an uncertainty for each of the daily mean values of 0.2π. The 18 daily mean values are found to vary by more than this estimated uncertainty and have a standard deviation of 0.4π. The source of this excess variability is not known at this time.

As a preliminary search for secular variability, the 18 daily mean values were fitted to a linear trend. The result is an apparent decrease in solar radius of 0.2 ± 0.2 π/year. Because the size of the uncertainty, this result is not statistically significant.

REFERENCES

1. Eddy, J. A.; and Boornazian, A. A.: Secular Decrease in the Solar Diameter, 1863-1953. Bull. Am. Astron. Soc. 11, 1979, p. 437.

2. Meyermann, B.: Zur Pulsation der Sonne, Astron. Nachr. 279, 1950, p. 45.

ESTIMATING SHORT-TERM SOLAR VARIATIONS BY A SIMPLE ENVELOPE MATCHING TECHNIQUE

Kwing L. Chan
Applied Research and Systems

Sabatino Sofia
NASA/Goddard Space Flight Center
Laboratory for Planetary Atmospheres

ABSTRACT

A simple matching technique is explained which allows the computation of the response of the solar surface to perturbations which occur at any depth within the convective envelope of the Sun. This technique was applied to a perturbation of the convective efficiency (α-mechanism), and of the non-gas component of the pressure (β-mechanism) in different regions of the convection zone. The results indicate that either perturbation affects the solar luminosity. However, the α-mechanism has little effect in the solar radius, regardless of the location of the perturbed region, whereas the β-mechanism produces radius changes that become quite large if the location of the perturbed region is deep within the solar convection zone.

Changing the mixing length ratio used in computing the solar convective envelope is one mechanism capable of producing a fluctuation of the solar luminosity (refs. 1,2). However, there are other possible mechanisms which can also produce solar variations. To facilitate the investigation of the effects of as many possible mechanisms as possible in an economical way, we shall describe here a simple matching technique for quick estimation of these effects. This technique cannot give the time evolution of the fluctuations, and the action of the perturbing mechanism must be confined to be above the bottom of the convection zone. However, this technique can give a rather accurate estimate of the amplitudes of the solar variations with very little computational effort. We shall first describe the mechanics of the technique, then its justification, and finally some applications.

The procedure to follow is very simple. First, one calculates the envelope structure of an unperturbed solar model using some standard stellar envelope code (cf. ref. 3). Then the perturbing mechanism must be implemented in the evelope code, and a new envelope matched to the old envelope at a mass point near the base, but within, the convection zone. By "matching", we mean that the luminosity and the outer radius of the new envelope are changed until its pressure, temperature, and radius at the matching mass point become the same as those of the old envelope. After the fitting parameters are determined, the differences between the new and old envelopes can be interpreted to be the variation of the solar luminosity and the spontaneous change in the solar radius.

To justify this proceedure, we need to answer the following questions:

1. There are three variables: pressure, temperature, and radius, to be matched but there are only two fitting parameters - luminosity, and outer radius. Is it possible to match all three variables consistently?

2. The matching point at which the matching is done is not a priori fixed. Are the results of matching insensitive to this point?

3. The envelope integrations need to assume the luminosity to be uniform above the matching point. Is this procedure valid?

4. The variation of the pressure, temperature, and radius at the matching point are taken to be zero. Is this a valid approximation? For how long?

To answer questions 1 and 2, let us examine Table 1. This table shows the results of perturbing the solar envelope by increasing the mixing length ratio α by +0.1% from an initial value of 1.5. The first row at a given mass point (M_r) gives the results obtained by matching the pressure and the temperature at M_r, and the second row gives the results obtained by matching the radius and the temperature at M_r. As one can see, the variations in luminosity are identical, and the variations in the solar radius all agree to within 15%. The ratios between δ log R and δ log L are all about 6×10^{-4}, in very good agreement with time-dependent calculations (refs. 4 and 5). Therefore, matching either pressure-temperature or radius-temperature give consistent results. Furthermore, this table shows that the results are insensitive to the exact location of the matching point.

For the luminosity L to be uniform in a perturbed envelope, it is necessary that the time scale for the relaxation of a non-uniform distribution of the luminosity be very short. To estimate this time scale, we derived a diffusive type equation for a non-uniform distribution of luminosity ΔL in the almost-adiabatic convective region (using mixing-length theory of convection)

$$\frac{\partial}{\partial t} \Delta L = \frac{\varepsilon}{\tau} \frac{\partial}{\partial \ln p} \left(\frac{1}{\varepsilon} \frac{\partial}{\partial \ln p} \Delta L \right), \qquad (1)$$

where $\varepsilon = 4\pi \rho C_p T r^2 H$, $\tau = (4/3\ \alpha)\ (H/V)$. H is the pressure scale height, V is the convective velocity, and all other symbols have their standard meaning. The important time scales here is τ which is close to the eddy-turn-over time scale of the convective turbulence. Near the bottom of the convection zone, this time scale is of the order of half a month. Therefore, convection can smooth the luminosity distribution in very short time. This smoothing time scale constitutes the lower limiting time scale for the applicability of our approach.

The upper limiting time scale for the validity of this approach can be estimated as follows. An extra energy leaving the solar surface $\delta L \Delta t$ is approximately given by half of the gravitational energy released by a collapse of a layer of mass ΔM by an amount δr in radius

$$\delta L \Delta t \approx \frac{1}{2} \left[\frac{G M_o}{R_o^2} \right] \left(\frac{\delta r}{r} \right) \left(\frac{\Delta M}{M_o} \right). \qquad (2)$$

This provides an estimate for the growth of the collapse δr in time. If we

want the estimate of the variation in the outer radius δR to be not completely washed away by the collapse δr, then we need δr to be less than δR. This implies a restriction on the time period given by the inequality

$$\Delta t < 2 \times 10^5 \, W \text{ (yrs)} \qquad (3)$$

where $W = \delta \log R / \delta \log L$. In the case for perturbing α, $W = 6 \times 10^{-4}$; therefore Δt must be less than 120 yrs.

We have answered all the previous questions on the validity of the approach. Now, let us briefly describe some results of applying the present technique to a calculation of the variations induced by a perturbation of the magnetic pressure in the upper convection zone. Let β be defined as P_m/P_t, where P_m is the material pressure, and P_t is the total pressure, which in addition to P_m, includes the radiation pressure and some "effective" magnetic pressure. Treated as a fluid, the magnetic field actually behaves quite similar to the radiation in the sense that its polytropic index is also 3. A lowering of β would mean that the fraction of electromagnetic pressure is raised relative to the material pressure. Figure 1 shows the results obtained by suddenly lowering β by 0.1% in the region between the solar surface and M_r (the horizontal axis). The solid curve shows $\delta R/R_o$ and the broken curve shows $\delta L/L_o$ as a function of the depth of the perturbation (the matching point for all cases is at $M_r = 0.99$). One can see that as the perturbation goes deeper, the effects become more prominent. Especially, the perturbation in radius increases substantially when $\delta\beta$ occurs below $M_r = 1\text{-}10^{-4} \, M_o$; the $\delta \log R/\delta \log L$ ratio becomes close to 0.1. However, we must stress that the present calculation is very preliminary, and it can only give an order of magnitude guide to the general response of the convective envelope to β perturbations.

REFERENCES

1. Ulrich, R. K. Solar Neutrinos and Variations in the Solar Luminosity. Science, 190, 619-624, 1975.

2. Sofia, S.; O'Keefe, J. R.; Lesh, J. R.; and Endal, A. S. Solar Constant: Constraints on Possible Variations Derived from Solar Diameter Measurements. Science, 204, 1306-1308, 1979.

3. Cox, J. P.; and Giuli, R. T. Principles in Stellar Structure (New York: Gordon and Breach), 1968.

4. Twigg, L. W.; and Endal, A. S. Thermal Perturbation of the Sun. Workshop on Variations of the Solar Constant. NASA CP-2191, 1981.

5. Sweigart, A. V. Effects of Changes in Convective Efficiency on the Solar Radius and Luminosity. Workshop on Variations of the Solar Radius and Luminosity. Workshop on Variations of the Solar Constant. NASA CP-2191, 1981.

TABLE 1. RESULTS FOR PERTURBING BY

$\Delta\alpha$ + 0.1% FROM α = 1.5

Matching Point M_r	δ log (L)	δ log (R)	$\frac{\delta \log (R)}{\delta \log (L)}$
.9850	.326(-3)	.189(-6)	.579(-3)
	.326(-3)	.193(-6)	.591(-3)
.9900	.326(-3)	.189(-6)	.580(-3)
	.326(-3)	.185(-6)	.566(-3)
.9990	.326(-3)	.182(-6)	.557(-3)
	.326(-3)	.170(-6)	.521(-3)
.9999	.326(-3)	.185(-6)	.567(-3)
	.326(-3)	.165(-6)	.507(-3)

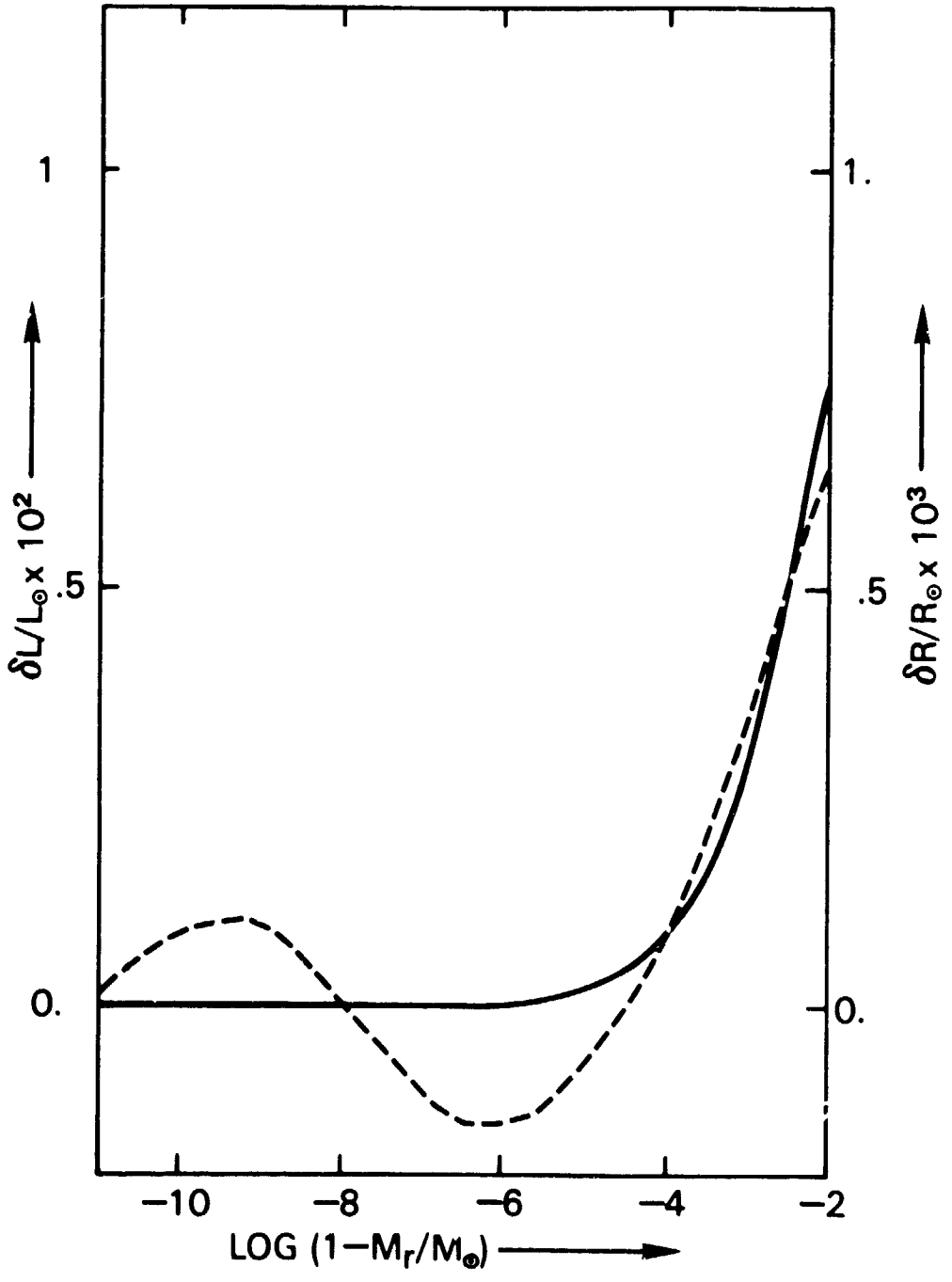


Figure 1. Variations induced by a sudden decrease in β (0.1%) in the region above M_r. Solid line shows δR, and broken line δL.

THERMAL PERTURBATION OF THE SUN*

L. W. Twigg and A. S. Endal
Department of Physics and Astronomy
Louisiana State University

ABSTRACT

An investigation of thermal perturbations of the solar convective zone via changes in the mixing length parameter have been carried out, with a view toward understanding the possible solar radius and luminosity changes that have been cited in the literature. The results show that: (a) a single perturbation of α is probably not the cause of the solar radius change given in ref. 2, and (b) the parameter $W \equiv d\ell nR_\odot/d\ell nL_\odot$ can not be characterized by a single value, as has been implied in recent work (refs. 2-5)

INTRODUCTION

Recent observations of possible changes of the solar luminosity (ref. 1) and radius (ref. 2) have spurred theoreticians to try to model the physical mechanisms that might produce such changes. One possible mechanism involves thermal perturbations of the solar convection zone. Such perturbations can be modeled (to first order) by perturbing the mixing length parameter α(equal to the ratio of the mixing length to the pressure scale height), used in the standard mixing length theory of convection. In this paper we present the results of such an analysis. Other work in this area can be found in references 2 to 5.

METHOD OF ANALYSIS

The stellar evolution code described by Endal and Sofia (ref. 6) was used for this investigation, utilizing the following assumptions and input data:

(1) Spherically symmetric Sun in hydrostatic equilibrium
(2) Latest Los Alamos opacities and equations of state (ref. 7)
(3) Nuclear reaction rates of Fowler, et al. (ref. 8)
(4) Standard mixing length theory of convection.
(5) The use of 700 interior zones and a small ($6\times10^{-8}M_\odot$) static envelope, to insure numerical accuracy.

*Research supported in part by NASA grant NAO5-13.

A one solar mass ZAMS model was evolved to the present age of the Sun ($T_\odot = 4.7 \times 10^9$ yrs.). A match to the present luminosity and radius of the Sun was found by varying the initial hydrogen abundance, X, and the value of α. The adopted values were X = 0.71242 and α = 2.21772. A sequence was then calculated such that the time step was slowly decreased until ΔT = 1 yr. at $T_\odot$. A perturbation of α in the range 0.05% to 4% was then introduced, and the full non-linear, time-dependent evolution followed either for 200 yrs. (with ΔT = 1 yr.) or for 6×10^7 yrs. (allowing the time step to increase), at which time the normal evolutionary effects dominated the solution. In order to test the effect of varying the size of the small time step, similar sequences were calculated using timesteps of 0.75, 2, and 5 yrs. No difference in the resulting models was seen. The final results are discussed in the next section.

RESULTS

The results of this analysis are summarized in Table I. We can divide the results into three areas, which are discussed below. In the following discussion we adopt the following definitions:

$$\Delta(x) \equiv x_{new} - x_{old} = \text{change in } x.$$

$$\delta(x) \equiv [\Delta(x)*100]/x_{old} = \text{percent change in } x.$$

(1) Solar Luminosity: As can be seen in Table I, the percentage change in $L_\odot$ is of the same sign as $\delta(\alpha)$, and of a much larger value than the corresponding $\delta(R_\odot)$ (here $R_\odot \equiv$ solar radius at optical depth 2/3). The relation between $\delta(L_\odot)$ and $\delta(\alpha)$ is linear for $0.75 \leq \delta(\alpha) \leq 4.00$, and can be written down as: $\delta(L_\odot) = 0.76 \cdot \delta(\alpha)$. The small change in slope at $\delta(\alpha) = 0.75$ will be discussed later. The characteristic time scale to recover its initial value is a thermal time scale of $\approx 6 \times 10^6$ yrs.

(2) Solar Radius: In the solar photosphere ($\tau \leq 2/3$), $\delta(r)$ shows an initial change of the same sign as $\delta(\alpha)$, followed by a subsequent relaxation toward the new equilibrium radius (larger for $\delta(\alpha)$ negative, smaller for $\delta(\alpha)$ positive) on a much more rapid time scale than $\delta(L_\odot)$. For $\delta(\alpha) = 2$, the time scale to cross its original value is ≈ 200 yrs. For $\tau > 2/3$, (r) simply shows an immediate relaxation toward the new equilibrium value. As Figure 1 shows, $\delta(R_\odot)$, and $\delta(L_\odot)$ to a lesser extent, changes its behavior with $\delta(\alpha)$ at $\delta(\alpha) = 0.75$. The reason for this can be traced to the finite interpolation scheme used for the envelope boundry conditions. Thus the values for $\delta(R_\odot)$ and $\delta(L_\odot)$ are probably not reliable below $\delta(\alpha) = 0.75$.

(3) W: A parameter used by investigators in the field to characterize the changes of radius and luminosity due to any perturbation is $W \equiv d\ell n\, R_\odot / d\ell n\, L_\odot = \delta(R_\odot)/\delta(L_\odot)$. Figure 2 shows a plot of W_o vs $\delta(\alpha)$ for the data in Table I. Here $W_o = W$ for the first model after the perturbation is introduced. As noted, the values of W are not reliable for $\delta(\alpha) < 0.75$. We see that (a) since $\delta(R_\odot)$ changes on a much faster time scale than $\delta(L_\odot)$, W will change with time (and will change sign, e.g. after $\approx$ 200 yrs. for $\delta(\alpha) = 2\%$), and (b) as Figure 2 shows, the value of W_o changes with $\delta(\alpha)$. Hence, there really is no single value for W, as has been implied by earlier investigations.

CONCLUSION

The main conclusions that can be drawn from this investigation are as follows:

(1) $\delta(L_\odot) = 0.76\ \delta(\alpha)$, while $\delta(R_\odot)$ displays a marked non-linear behavior with $\delta(\alpha)$.

(2) There is no one value of W which generally characterizes thermal perturbations of the solar convective zone.

(3) Even for $\delta(\alpha) = 5$, $\delta(R_\odot) \approx 2.10^{-3}$ while $\delta(L_\odot) \approx 3$! Thus for $\delta(R_\odot) = 2.5 \times 10^{-2}$ (ref. 2), the extremely large $\delta(L_o)$ implied shows that a single perturbation of α is probably not the cause of the radius change, although further work for $\delta(\alpha) < 0.75$, and for a series of random α perturbations is definitely indicated.

(4) Since our results for $\delta(\alpha) < 0.75$ are probably incorrect, a value of W near 0.075 for small $\delta(\alpha)$ is not precluded. Thus our results show that by taking into account the full time-dependent, non-linear behavior of the problem, the entire range of W quoted in the literature (refs. 2 to 5) may be generated.

REFERENCES

1. Willson, R. C., Hudson, H. S., and Chapman, G. A., Observations of Solar Irradiance, 1981, Science, 211, 700.

2. Sofia, S., O'Keefe, J., Lesh, J. R., and Endal, A. S., Solar Constant: Constraints on Possible Variations Derived from Solar Diameter Measurements, 1979, Science, 207, 1306.

3. Spiegal, E. A., and Weiss, N. O., Magnetic Activity and Variations in Solar Luminosity, 1980, Nature, 287, 616.

4. Dearborn, D. S. P., and Blake, J. B., Is the Sun Constant?, 1980, Ap. J. 237, 616.

5. Gilliland, R. L., Solar Luminosity Variations, 1980, Nature, 286, 838.

6. Endal, A. S., and Sofia, S., Rotational History of the Sun: Spin Down of the Interior by Circulation Currents and Fluid Instabilities, 1981, Ap. J., in press.

7. Huebner, W. F., Merts, A. L., Magee, N. H. Jr., and Argo, M. F., Astrophysical Opacity Library, Los Alamos Scientific Laboratory Manual LA-6760-M, 1977.

8. Fowler, W. A., Caughlan, G. R., and Zimmerman, B. A., Thermonuclear Reaction Rates, II, 1975, Annual Reviews of Astronomy and Astrophysics, 13, 69.

Table 1

$\delta\alpha$	$\delta R_\odot$	$\delta L_\odot$	$W = \frac{\delta R_\odot}{\delta L_\odot}$
.05	3.818×10^{-5}	.0379	1.007×10^{-3}
.1	7.644×10^{-5}	.0758	1.008
.3	2.303×10^{-4}	.2277	1.015
.75	5.802×10^{-4}	.5701	1.018
.8	5.974×10^{-4}	.6055	.987
.85	6.146×10^{-4}	.6410	.959
.9	6.319×10^{-4}	.6764	.934
.95	6.493×10^{-4}	.7119	.912
1.0	6.667×10^{-4}	.7473	.892
1.25	7.577×10^{-4}	.9256	.819
1.5	8.460×10^{-4}	1.103	.767
2.0	10.262×10^{-4}	1.459	.703
2.5	12.119×10^{-4}	1.815	.668
3.0	14.033×10^{-4}	2.173	.646
3.5	16.038×10^{-4}	2.533	.633
4.0	18.047×10^{-4}	2.893	.624

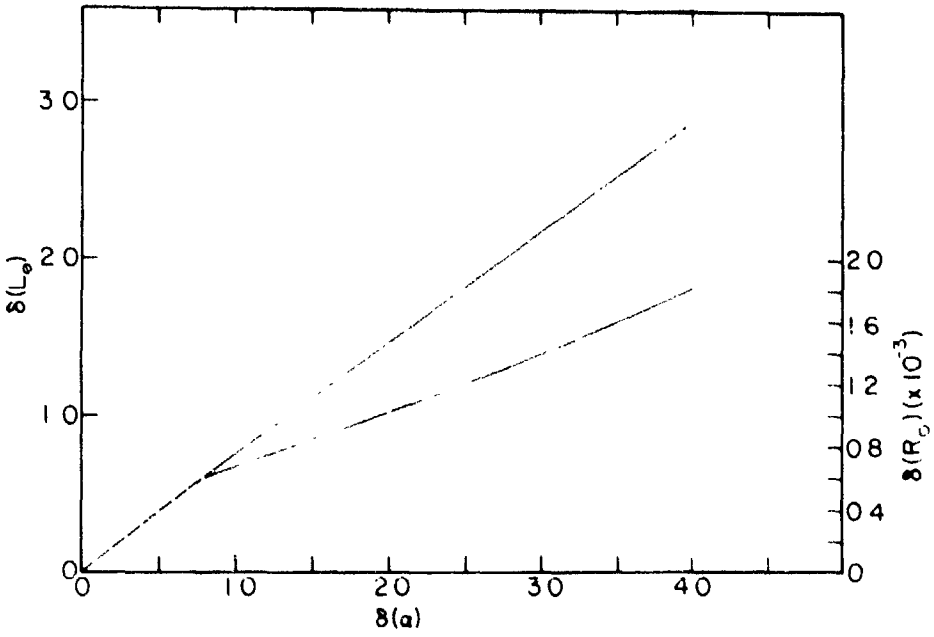


Figure 1. The relation between $\delta(L_\odot)$ and $\delta(a)$, (solid line) and between $\delta(R_\odot)$ and $\delta(a)$, (interrupted line). The data for $\delta(a) < 0.75$ is unreliable (dashed line).

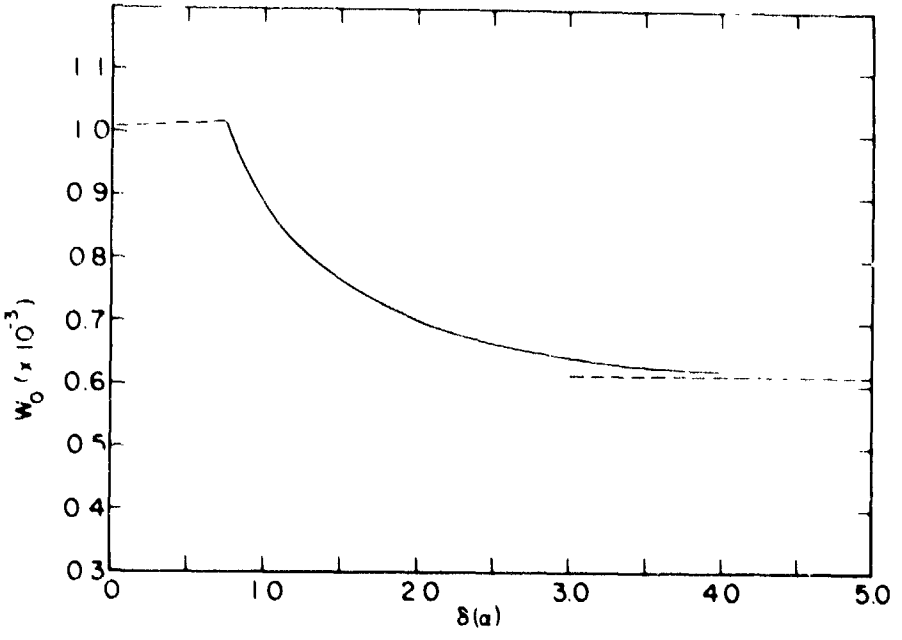


Figure 2. The relation between $W_\odot$ and $\delta(a)$. The data for $\delta(a) < 0.75$ is unreliable (dashed line). The dashed line for $a > 3$ is only intended as a guide for the eye.

D/6

N82 17030

EFFECTS OF CHANGES IN CONVECTIVE EFFICIENCY ON THE SOLAR RADIUS AND LUMINOSITY

A. V. Sweigart
NASA/Goddard Space Flight Center

ABSTRACT

A sequence of solar models has been constructed in order to investigate the sensitivity of the solar radius and luminosity to small changes in the ratio α of the mixing length ℓ to the pressure-scale height H_p throughout the solar convective envelope. The basic procedure for determining this sensitivity was to impose a perturbation in α within the convective envelope and then to follow the resulting changes in the solar radius ΔR and luminosity ΔL for the next 10^6 yrs. These calculations gave the following results. 1) A perturbation in α produces immediate changes in the solar radius and luminosity. Initially ΔL and $\Delta\alpha$ are related by $\Delta L/L = 0.30\Delta\alpha/\alpha$. 2) The value of the ratio $W = \Delta\log R/\Delta\log L$ is strongly time dependent. Its value just after the perturbation in α is 6.5×10^{-4}. 3) The ratio $H = (\Delta\log L)^{-1} d\,\Delta\log R/dt$ is much less time dependent and is a more suitable means for relating the changes in the solar radius and luminosity. 4) Both of these ratios imply that for any reasonable change in the solar luminosity the corresponding change in the solar radius is negligible.

I. INTRODUCTION

During this workshop there has been much discussion about possible changes in the solar radius and luminosity over timescales ranging from a year or less to a few hundred years. Because of the keen interest in this topic and because of its obvious relevance to climatic conditions here on earth it is of considerable importance to determine the sensitivity of the solar radius and luminosity to changes in the interior structure of the Sun. Knowledge of this sensitivity together with observational data on any radius and luminosity changes would greatly help in understanding the characteristics of the physical processes operating within the solar interior and, as a result, in understanding the influence which these processes might have on the Sun's future behavior. In addition, it is of considerable importance to determine theoretically the relationship between changes in the solar radius and luminosity resulting from interior perturbations, since then observational data on one of these changes could be used to estimate the size of the other (ref. 1), provided, of course, that the physical process causing the perturbations has been properly identified.

There are many ways in which the interior structure of the Sun might be perturbed. In approximately the outer 2 per cent of the Sun's mass the outward energy flux is carried largely by convection. Since convection in

these layers is a turbulent process, it is entirely plausible that there may be random fluctuations in the efficiency of energy transport due, for example, to statistical fluctuations in the number of convective cells or to changes in the flow pattern. The convective envelope therefore represents one part of the Sun where interior perturbations might be expected naturally to arise.

Interior perturbations might also occur within the radiative core. Such perturbations would alter both the thermal and hydrostatic structure of the Sun. The thermal readjustment induced by such perturbations would take place over a Kelvin timescale which for the entire Sun is about 10^7 yrs (ref. 2). The hydrostatic readjustment, however, would take place on a dynamical timescale which is on the order of minutes (ref. 2) and would therefore manifest itself almost instantaneously as a change in the solar radius.

Studies of other stars provide some evidence that observable changes in the radius can result from perturbations within the core. The pulsation period of a class of variable stars known as RR Lyrae stars, found both in globular clusters and in the field, can be accurately determined by using observations spanning several decades. It has been found that the pulsation periods of the RR Lyrae stars typically vary at the rate of a few parts in 10^5 per century. Such changes in the pulsation period can be readily interpreted as changes in the mean stellar radius from one pulsation period to the next. The observed rates of period change considerably exceed the values expected from the normal evolution of the RR Lyrae stars - a fact that has proved to be a long-standing problem. Recent theoretical studies of RR Lyrae models (ref. 3) have shown that perturbations within the core of these stars can reproduce the observed characteristics of the period changes and can thus offer a reasonable solution for this problem. This result suggests that the radiative core of the Sun may also be a likely site for the perturbations responsible for any changes in the solar radius and luminosity.

The objective of the present paper is to give the results of one way of perturbing the solar interior, namely, by changing the efficiency of energy transport by convection throughout the convective envelope. In computing the structure of the solar convective envelope it is necessary to know the value of the convective-temperature gradient, i.e., the actual temperature gradient, at each point. The value of this gradient is determined by the requirement that the total energy flux carried by both convection and radiation be equal to the actual outward energy flux. The convective gradient can range between two limiting values, namely, the adiabatic- and the radiative-temperature gradients, depending on the degree of convective efficiency. When convection is very efficient, the convective gradient approaches the adiabatic gradient. This is normally the situation at higher densities and temperatures when the thermal energy content of the convective cells is relatively large. At lower densities and temperatures, convection can become quite inefficient, and, as a result, the convective gradient becomes significantly superadiabatic and can in fact approach the radiative gradient, which is defined to be the temperature gradient that would exist if all of the outward energy flux were carried by radiation.

The calculation of the convective gradient as a function of the physical conditions at each point in the convective envelope is generally done according to the prescription of the mixing-length theory. In this theory the turbulent convective motions which actually cover a wide range of scale lengths are assumed to be represented by convective cells that travel a characteristic length ℓ before dissolving into the surrounding medium. The mixing length ℓ is the main parameter governing the convective efficiency. An increase in ℓ enhances the convective efficiency, thereby lowering the convective gradient. Conversely, a decrease in ℓ reduces the convective efficiency, since the convective cells then cannot transport their excess thermal energy as far before dissipation. Ordinarily the value of ℓ at each point is expressed in terms of some scale height such as the pressure-scale height H_p (= dr/dlnP). In this paper we will study the consequences of changing the ratio α (= ℓ/H_p) and hence the convective efficiency in the solar convective envelope. Two points should, however, be kept in mind when considering the following results. First, there are other ways in which the properties of the solar convection could be altered, and hence this paper examines only one type of convective perturbation. Secondly, the perturbation results assume that the mixing-length theory adequately determines the structure of the solar convective envelope at least as far as small perturbations away from the equilibrium structure are concerned.

In the next section we describe first the unperturbed structure of the solar convective envelope and then the effects which a perturbation in α has on this structure at various times following the perturbation. The changes in the solar radius and luminosity resulting from a perturbation in α and the relationship between these changes are discussed in sections III and IV, respectively. We emphasize in section IV the advantages of using the time rate of change of the radius perturbation rather than the radius perturbation itself when relating the radius and luminosity perturbations. Finally, a summary of the main points is provided in section V.

II. SOLAR CONVECTIVE ENVELOPE

UNPERTURBED STRUCTURE

In order to examine the unperturbed structure of the solar convective envelope, one must first obtain a solar model with the proper luminosity and radius at an age of 4.7×10^9 yrs following the zero-age main-sequence (ZAMS) phase. The properties of a solar model are dependent on the assumed composition, i.e., the helium abundance Y and the heavy-element abundance Z, and on α. For the present calculations Z was taken to be 0.02. The luminosity of a solar model is particularly sensitive to Y, since changes in Y affect the mean molecular weight and hence the hydrostatic structure, leading to a change in the central temperature. This in turn alters the rate of hydrogen burning due to the strong temperature dependence of the nuclear reaction rates. On the other hand, α primarily affects the convective envelope and thus the radius. Several trial sequences showed that the values

$$Y = 0.2317, \alpha = 1.70 \quad (1)$$

reproduce the present solar luminosity and radius with an error of about 0.2 per cent. Accordingly these values were adopted for the model computations, and a standard evolutionary sequence was then computed from the ZAMS phase to the present Sun. The ZAMS luminosity and radius of this solar sequence were 0.723 $L_\odot$ and 0.893 $R_\odot$.

The unperturbed structure of the convective envelope in the present Sun is perhaps best illustrated by the behavior of the adiabatic-, convective- and radiative-temperature gradients (= dlogT/dlogP), denoted by ∇_a, ∇_c and ∇_r, respectively. These three gradients are plotted in Figure 1 as functions of the logarithm of the amount of mass between the surface and the given point. Here M_r is the amount of mass within a distance r from the center of the Sun. The convective envelope in this model contains 0.016 $M_\odot$, corresponding to $\log (M_\odot - M_r) = -1.796$. For a fully ionized nondegenerate gas with negligible radiation pressure ∇_a equals 0.40, and we note that ∇_a approaches this value throughout the inner part of the convective envelope, i.e., for $\log (M_\odot - M_r) > \sim -4$. Between $\log (M_\odot - M_r) = -11$ and -4, ∇_a is depressed due to the ionization of hydrogen and the first and second ionizations of helium. The difference between ∇_c and ∇_a is very small for $\log (M_\odot - M_r) > \sim -7$. This "adiabatic" region contains the bulk of the mass within the convective envelope. In this region energy transport by convection is very efficient with radiation making only a negligible contribution to the outward flux, since $\nabla_r >> \nabla_c$. Because of this high convective efficiency ∇_c will not be very sensitive to α. Just the opposite is true in the layers near the surface ($\log (M_\odot - M_r) < \sim -9$), where the convection becomes strongly superadiabatic. As one goes outward through this "superadiabatic" region, ∇_c begins to exceed ∇_a substantially with the maximum superadiabaticity being reached at $\log (M_\odot - M_r) = -10.5$. The convective efficiency in the superadiabatic region is low due to the low density and thermal energy content of the convective cells. In the layers nearest the surface ∇_c approaches ∇_r and hence the energy transport there is largely by radiation. The structure of the superadiabatic region will be strongly dependent on α. The transition between the adiabatic and superadiabatic regions occurs around $\log (M_\odot - M_r) = -8$.

PERTURBED STRUCTURE

Before discussing the quantitative results from detailed solar model computations it is worthwhile to mention first some further features of the solar convective envelope and to consider the physical reasons for the way in which the Sun responds to a change in α. The superadiabatic region contains little mass and has only a small thermal energy content. As a result, the thermal timescale of the region is quite short, on the order of 1 day (ref. 4), and consequently the superadiabatic region rapidly readjusts to any change in α. Within the adiabatic region ∇_c is nearly independent of α, and therefore changes in α that are confined to this region will not significantly affect

the structure of the convective envelope. However, a change in α throughout the entire convective envelope will alter the boundary conditions at the top of the adiabatic region, and this in turn will force the adiabatic region to undergo both a dynamical and thermal readjustment. The dynamical response of the adiabatic region will restore hydrostatic equilibrium on a timescale of minutes and will thus be practically instantaneous. The contraction (or expansion) associated with this dynamical readjustment will release (or absorb) gravitational potential energy, thereby perturbing the outward energy flux L_r and causing the adiabatic region to depart from thermal equilibrium. The timescale for restoring thermal equilibrium is on the order of 10^5 yrs (ref. 5). The key point to remember is that the response of L_r and hence the solar luminosity is set, not by the thermal timescale of the adiabatic region, but by the much shorter thermal timescale of the superadiabatic region. Therefore one would expect a change in α to show up almost immediately as a perturbation in the solar luminosity.

Let us now outline the sequence of events to be expected if, for example, α increases. After about 1 day the superadiabatic region will have readjusted both thermally and hydrostatically. As is well-known from stellar model computations, an increase in α leads to a contraction of the adiabatic region and hence to the concomitant release of gravitational potential energy, resulting in an increase in the outward flux L_r and thus in the solar luminosity. The superadiabatic region will then expand in order to carry the additional outward flux, since this is the normal reaction of a region in which energy transport by radiation is important (ref. 4). Thus one has a situation in which the bulk of the convective envelope contracts on a timescale of $\sim 10^5$ yrs while the outermost layers initially expand on a timescale of days. Observationally this would appear as a sudden increase in the solar radius followed by a gradual decrease. A similar sequence of events would also occur if α decreases except that all of the perturbations would have opposite signs.

In order to verify the above predictions quantitatively, a sequence of solar models was constructed in which the time step between models, which is normally set by the nuclear timescale of the core, was gradually reduced to 1 yr. This choice for the minimum time step was made in order to follow the rapid changes expected in the solar radius and luminosity while avoiding the numerical difficulties sometimes encountered when even shorter time steps are used. At this point in the calculations the value of α was increased by $\Delta\alpha = 0.01$ throughout the convective envelope, and the subsequent evolution of the perturbed solar models was followed for about the next 10^6 yrs. After the change in α the time step was slowly increased but was always small compared with the timescale on which the perturbations were changing. The size of the perturbations resulting from this change in α are very small compared with the numerical accuracy of typical solar models. For this reason it was essential to maintain a high degree of numerical accuracy and especially to minimize the importance of numerical noise during the computations.

When constructing a stellar model one usually treats the outermost layers differently from the interior. In the outermost layers the stellar structure equations are integrated inward from the surface to some interior fitting point under the assumption of constant L_r. This is equivalent to ignoring any changes in the gravitational potential energy, i.e., to assuming thermal equilibrium. Given several of these integrations, one can then define the outer boundary conditions needed for the interior solution. Inside the fitting point the stellar structure equations are replaced by difference equations which are then solved by an iterative procedure. In the present solar models the fitting point was located at $\log (M_\odot - M_r) = -6$. Since according to the previous discussion the thermal timescale of the outermost 10^{-6} $M_\odot$ of the Sun is very short, our implicit assumption of thermal equilibrium in these layers should be justified. About 175 integration steps based on a high-order predictor-corrector procedure were used in computing the layers above the fitting point. Interior to the fitting point there were 247 mesh points of which 88 were in the convective envelope.

There are many sources of numerical noise which can enter into solar model computations. For example, stellar structure programs frequently contain iterative procedures for determining the density from the equation of state, the degree of ionization from the Saha equations and the superadiabaticity within convective regions. Tight convergence of these iterative procedures as well as the iterative procedure involved in the overall convergence of the models was required at all times. In addition, no changes were permitted in either the number or distribution of the mesh points. Such changes in the mesh points could introduce spurious perturbations by altering the truncation error with which the difference equations represent the basic differential equations of stellar structure. Special attention must therefore be paid to these as well as a number of other sources of numerical noise if reliable results are to be obtained. To insure that numerical noise was not important in the present calculations, we constructed an additional solar sequence in which the perturbation in α was a factor of 10 greater, i.e., $\Delta\alpha = 0.10$. The only difference was the expected scaling in the size of the perturbations by a factor of 10. In particular, the ratio of the perturbations in the solar radius and luminosity changed by less than 2 per cent.

Let us now consider some of the quantitative results for the readjustment of the solar convective envelope after the perturbation $\Delta\alpha = 0.01$. Figure 2 illustrates the difference in the radius $\Delta \log r$ between a perturbed model and the basic unperturbed model as a function of $M_\odot - M_r$ within the convective envelope. The four curves labelled a, b, c and d correspond to four perturbed models having ages of 1, 4900, 49,000 and 310,000 yrs, respectively, following the perturbation in α. The contraction is not yet noticeable in model a, because the time elapsed since the perturbation has been too short. Moreover, the increase in the surface radius $\Delta \log R$ in model a due to the expansion of the superadiabatic region amounted to only 5×10^{-7}. By model d the rate of contraction has slowed substantially so that this model is approaching the equilibrium structure for the new value of α. The amount of the contraction

is considerably greater nearer the surface, and thus the convective envelope does not contract uniformly. This contraction increases the weight of the convective envelope on the radiative core, thereby causing the core also to contract, as indicated in Figure 2 for $M_\odot - M_r > 0.016\ M_\odot$.

The rate of release of gravitational potential energy ε_g in ergs/gm/sec within the convective envelope is shown in Figure 3 for each of the four perturbed models plotted in Figure 2. The maximum rate of contraction of the convective envelope occurs immediately after the perturbation in α, and thus the largest values of ε_g are produced at this time. However, the radiative core does not begin to contract until after there has been a decrease in the radius of the convective envelope and hence a change in the boundary conditions at the edge of the core. This explains why the release of gravitational energy in the core is negligible in model a while it becomes important in the later models. We note that ε_g is negative for $M_\odot - M_r < 6 \times 10^{-4}\ M_\odot$ in model a due to the expansion of the outer layers of the convective envelope. In the present calculations this expansion disappears 1 year after the perturbation in α; it might actually disappear sooner if shorter time steps are used. The slowing-down of the contraction with time, as indicated by the decrease in ε_g, is apparent in going from models a to d.

The release of gravitational potential energy perturbs the outward flux L_r at each point within the convective envelope. This flux perturbation ΔL_r is illustrated in Figure 4, where the difference in L_r between each of the four perturbed models in Figure 2 and the unperturbed model is shown over the same interval in $M_\odot - M_r$ as in Figure 3. The behavior of the flux perturbation in time is somewhat complicated in the inner half of the convective envelope due to two competing effects. Between models a and c the contribution to the flux perturbation from the contraction of the core increases, while at the same time the contribution of the convective envelope decreases. The drop in ΔL_r for $M_\odot - M_r < 6 \times 10^{-4}\ M_\odot$ in model a is again associated with the initial expansion of the outermost layers.

The above discussion has focused on the structural readjustment that takes place within the solar convective envelope following a perturbation in α. We now turn our attention to the question of what potentially observable changes a perturbation in α might produce in the solar radius and luminosity.

III. CHANGES IN THE SOLAR RADIUS AND LUMINOSITY

The changes in the solar radius $\Delta \log R$ and luminosity $\Delta \log L$ during the first 8×10^5 yrs after the perturbation in α are presented in Figure 5. The zero-point of the time scale in Figure 5 as well as in all subsequent figures corresponds to the time t when the perturbation $\Delta\alpha = 0.01$ was imposed within the convective envelope. The response of the solar luminosity to this perturbation appears to be nearly instantaneous for the time resolution of this figure. Following the large initial response $\Delta \log L$ decays with an e-folding time on the order of a few times 10^5 yrs. By the latest times shown in Figure

5 the perturbed solar models are approaching their new equilibrium structure which is characterized by a decrease in log R and an increase in log L. The present results demonstrate that the initial response of the solar luminosity considerably exceeds the difference in log L between the unperturbed and new equilibrium states. Also plotted in Figure 5 is the change in the rate of hydrogen burning $\Delta \log L_H$. The contraction of the core, as indicated previously, raises the temperature in the layers near the center, thus increasing the rate of the nuclear reactions.

It is of some interest to examine the behavior of $\Delta \log L$ and $\Delta \log R$ immediately following the perturbation in α. Figure 6 shows this behavior for $\Delta \log L$. The time scale in this figure has been expanded by approximately a factor of 2000 compared with Figure 5 and consequently covers only the first 400 yrs after the perturbation in α. Even on this expanded timescale there is a sudden response of the solar luminosity at $t = 0$. This response would actually have been more abrupt if time steps less than 1 year had been used in the computations. This result confirms our previous conjecture that changes in convective efficiency of the type considered here will almost immediately affect the surface luminosity. We note from Figure 6 that $\Delta \log L$ is nearly constant over a timescale of several hundred years. From these results it follows that the change in the solar luminosity produced by a perturbation $\Delta\alpha$ is given by

$$\frac{\Delta L}{L} = 0.30 \frac{\Delta \alpha}{\alpha} \qquad (2)$$

for short times after the perturbation. A similar expression has been derived by Dearborn and Blake (ref. 4), who found a coefficient of 0.64 on the right-hand side of equation (2).

The more complicated behavior of $\Delta \log R$ is illustrated in Figure 7 for the same time interval as in Figure 6. The ordinate in Figure 7 has been expanded by roughly a factor of 1000 in comparison with Figure 5. The sudden increase of the solar radius due to the expansion of the superadiabatic region is readily apparent at $t = 0$. This initial expansion is followed by an overall contraction of the convective envelope and hence in the solar radius as the adiabatic region reacts to the change in α. At $t = 250$ yrs the radius again equals its unperturbed value. The maximum value of $\Delta \log R$ just after the perturbation in α was quite small, only 5×10^{-7}, which explains why the initial expansion was not evident in Figure 5. This maximum value of $\Delta \log R$ is related to the perturbation $\Delta\alpha$ by the equation

$$\frac{\Delta R}{R} = 2.0 \times 10^{-4} \frac{\Delta \alpha}{\alpha} . \qquad (3)$$

Two features of Figure 7 should be emphasized. First, the value of $\Delta \log R$ is strongly time dependent even over the short time interval covered by this figure. Second, the rate of change of $\Delta \log R$, $d \Delta \log R/dt$, is, in contrast, nearly constant.

In this section we have discussed each of the changes $\Delta \log L$ and $\Delta \log R$ separately. We now wish to consider how these changes are related to each other.

IV. RELATIONSHIP BETWEEN CHANGES IN THE SOLAR RADIUS AND LUMINOSITY

One of the objectives of previous studies (refs. 1, 4, 6, 7, 8) on the effects of perturbations in α was to determine the ratio

$$W = \frac{\Delta \log R}{\Delta \log L} . \tag{4}$$

This ratio can be straightforwardly obtained from the present calculations to give the results shown in Figure 8, where the time interval is the same as in Figures 6 and 7. The strong time dependence of W is immediately evident. In fact, the value of W changes sign at t = 250 yrs. Since $\Delta \log L$ is nearly constant over the time interval in Figure 8, this time dependence is actually a reflection of the strong time variation of $\Delta \log R$. The values of W in Figure 8 can be approximated by the equation

$$W(t) = 6.5 \times 10^{-4} - 2.5 \times 10^{-6}\, t, \tag{5}$$

where t is in years. The original estimates of $W(t = 0)$ ranged from 0.075 to 5×10^{-3} (refs. 1, 4). More recent determinations have averaged from 5×10^{-4} to 10×10^{-4} (refs. 6, 7, 8) and are therefore in agreement with the present value.

One would like to use W to determine, for example, the change in the solar luminosity associated with observational estimates for changes in the solar radius. However, there are two major disadvantages with using W for this purpose. First, it is only appropriate to use W if the perturbation in α has occurred during the time interval spanned by the radius observations. Otherwise any observed change in log R would actually be the change between two perturbed states rather than between the unperturbed and perturbed states. From the last section we know that a perturbation in α gives rise to changes $\Delta \log R$ and $\Delta \log L$ that persist for several times 10^5 yrs. Thus, if an observed change in the solar radius is ascribed to a perturbation in α, the probability that this perturbation occurred during the interval of the observations is very small. Second, there is the problem caused by the strong time dependence of W. Even if the first disadvantage is ignored, one must still know how much time has elapsed since the perturbation in α in order to compute the proper value of W from equation (5).

The above difficulties can be overcome by using an alternative expression relating $\Delta \log R$ and $\Delta \log L$, namely, the ratio

$$H = \frac{1}{\Delta \log L} \frac{d\, \Delta \log R}{dt} = \frac{1}{\Delta \log L} \frac{d \log R}{dt} . \tag{6}$$

This ratio is plotted in Figure 9, where the time interval is again the same as in Figures 6 and 7. The average value of H in Figure 9 is

$$H = -2.6 \times 10^{-6}\ \mathrm{yr}^{-1}. \qquad (7)$$

The variation in the value of H is substantially less than was the case for W. This result is not surprising in view of our previous comments that $\Delta \log L$ and $d\,\Delta \log R/dt$ in Figures 6 and 7 are nearly constant. The fact that H is not strongly time dependent can also be given a straightforward physical explanation. The rate of contraction of the convective envelope $d\,\Delta \log R/dt$ determines the rate of release of gravitational potential energy which in turn determines the luminosity change $\Delta \log L$. Thus $d\,\Delta \log R/dt$ and $\Delta \log L$ actually represent different ways of measuring the same quantity, namely, the mean value of ε_g in the convective envelope, and consequently we would expect this ratio to be approximately constant at least for short times following the perturbation in α. Over much longer time intervals, however, the value of H will change significantly, as is illustrated by Figure 10, but even here the relative change is much less than that shown by W in Figure 8. For example, after 10^5 yrs the value of H differs by only a factor of 2 from its value at $t = 0$. When using H to relate $\Delta \log L$ to an observed radius change, one is implicitly assuming that the perturbation in α occurred prior to the time of the observations, but, as mentioned before, this is very likely to be the case. We conclude therefore that the inherent disadvantages of the ratio W can be circumvented to a large extent by using the ratio H.

Dunham et al. (ref. 9) have reported a decrease in the solar radius of 0.70 ± 0.12 between 1925 and 1980 from measurements of the size of the path of totality during a number of solar eclipses. The corresponding change in log R is thus -3×10^{-4}. Let us now see what this observational result implies for the change in the solar luminosity under the assumption that a perturbation in α is responsible for the radius change. There are two cases to consider. First, let us assume that the perturbation in α occurred sometime after 1925 so that W is the appropriate ratio to use. From equation (5) it follows that $5.1 \times 10^{-4} \leq W \leq 6.5 \times 10^{-4}$. The change in log L determined from these values of W lies in the range $-0.62 \leq \Delta \log L \leq -0.49$, implying that the solar luminosity in 1925 differed from the present luminosity by a factor of 3 or 4. As the second case, let us assume that the perturbation in α occurred before 1925 so that we must apply the ratio H. The radius measurements then give $-6 \times 10^{-6}\ \mathrm{yr}^{-1}$ for the average value of $d\,\Delta \log R/dt$ since 1925. By combining this observational result with the value of H from equation (7), we find that $\Delta \log L = 2.2$, again implying an impossibly large change in the solar luminosity. The change in log L would have been even greater if a smaller value of $|H|$ had been used, as would be appropriate for later times according to Figure 10. We conclude therefore that the change in the solar radius since 1925 either has not been as large as reported by Dunham et al. or has been produced by some process other than the one studied in this paper.

V. SUMMARY

From the present results it is possible to draw the following conclusions:

1) Changes in the efficiency of convection throughout the solar convective envelope lead to sudden changes in both the solar radius and luminosity. The relationship between the change in the luminosity and the change in α is given by equation (2).

2) The value of the ratio $W = \Delta \log R / \Delta \log L$ is strongly time dependent. For this and other reasons W does not seem to be a very suitable means for relating changes in the solar radius and luminosity. Immediately after a perturbation in α the value of W is 6.5×10^{-4}.

3) A more satisfactory way to relate the radius and luminosity changes is represented by the ratio $H = (\Delta \log L)^{-1} \, d\, \Delta \log R/dt$. This ratio is much less time dependent, varying from -2.6×10^{-6} to -1.3×10^{-6} yr^{-1} during the first 10^5 yrs following a perturbation in α.

4) According to the present values of W and H, any observationally detectable change in the solar radius would imply an impossibly large change in the solar luminosity. Consequently changes in convective efficiency of the type considered here cannot be responsible for any observed radius changes in the Sun.

REFERENCES

1. Sofia, S., O'Keefe, J., Lesh, J. R., and Endal, A. S.: Solar Constant: Constraints on Possible Variations Derived from Solar Diameter Measurements, Science, vol. 207, 1979, p. 1306.

2. Schwarzschild, M.: Structure and Evolution of the Stars. Princeton University Press, 1958.

3. Sweigart, A., and Renzini, A.: Semiconvection and Period Changes in RR Lyrae Stars. Astron. and Astrophys., vol. 71, 1979, p. 66.

4. Dearborn, D. S. P., and Blake, J. B.: Is the Sun Constant? Astrophys. J., vol. 237, 1980, p. 616.

5. Dearborn, D. S. P., and Newman, M. J.: Efficiency of Convection and Time Variation of the Solar Constant. Science, vol. 201, 1978, p. 150.

6. Gilliland, R. L.: Solar Luminosity Variations, Nature, vol. 286, 1980, p. 838.

7. Sofia, S., and Chan, K. L.: Estimating Short Term Solar Variations by a Simple Envelope Matching Technique. Workshop on Variations of the Solar Constant. NASA CP-2191, 1981.

8. Twigg, L. W., and Endal, A. S.: Thermal Perturbation of the Sun. Workshop on Variations of the Solar Constant, NASA CP-2191, 1981.

9. Dunham, D. W., Dunham, J. B., Fiala, A. D., and Sofia, S.: Eclipse Radius Measurements. Workshop on Variations of the Solar Constant. NASA CP-2191, 1981.

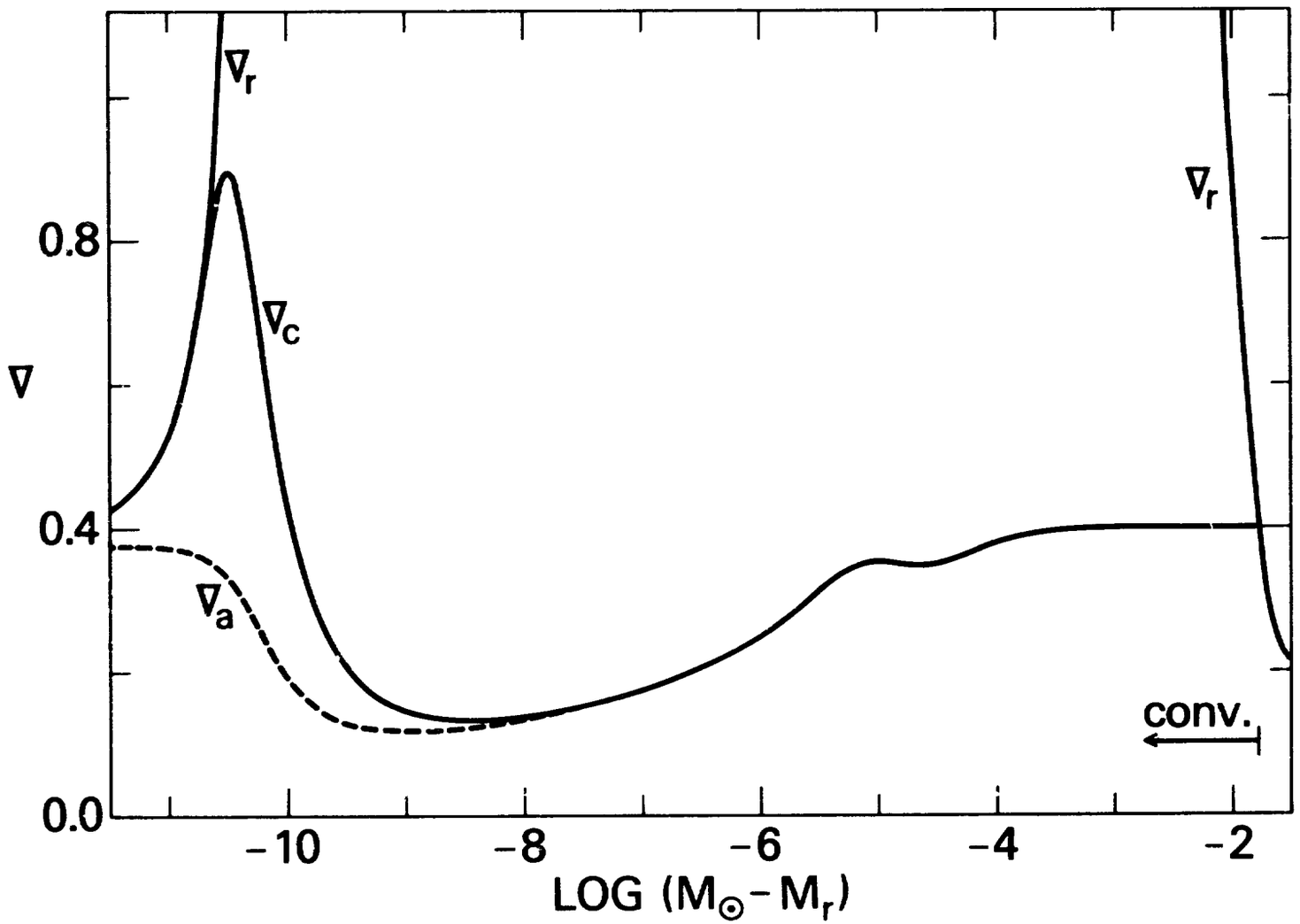


Figure 1. Variation of the adiabatic-, convective- and radiative-temperature gradients ∇_a, ∇_c and ∇_r, respectively, as functions of log ($M_\odot$ - M_r) within the solar convective envelope

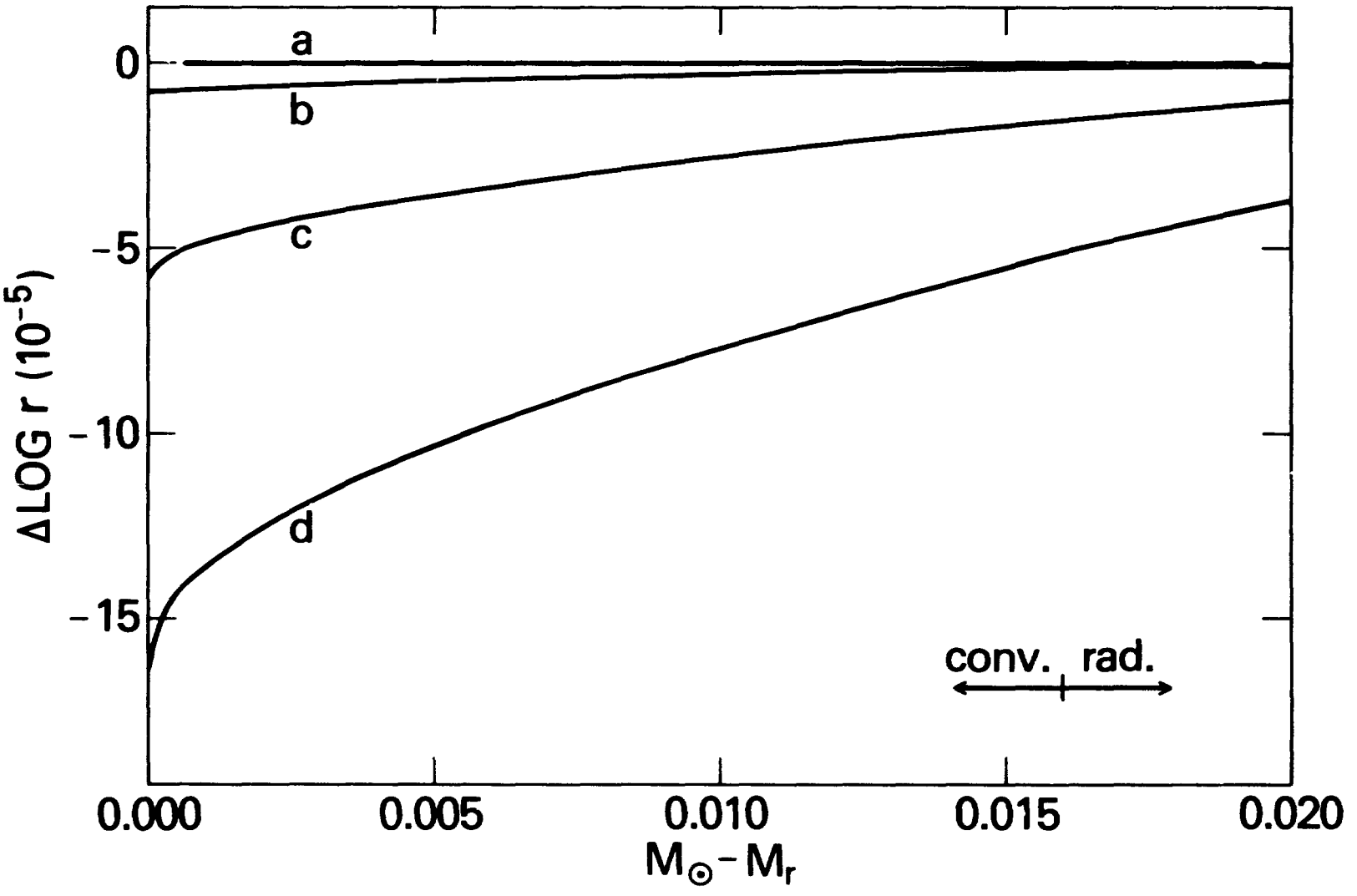


Figure 2. Perturbation in the radius Δlog r as a function of $M_\odot - M_r$ within the solar convective envelope at four times following the perturbation $\Delta\alpha = 0.01$

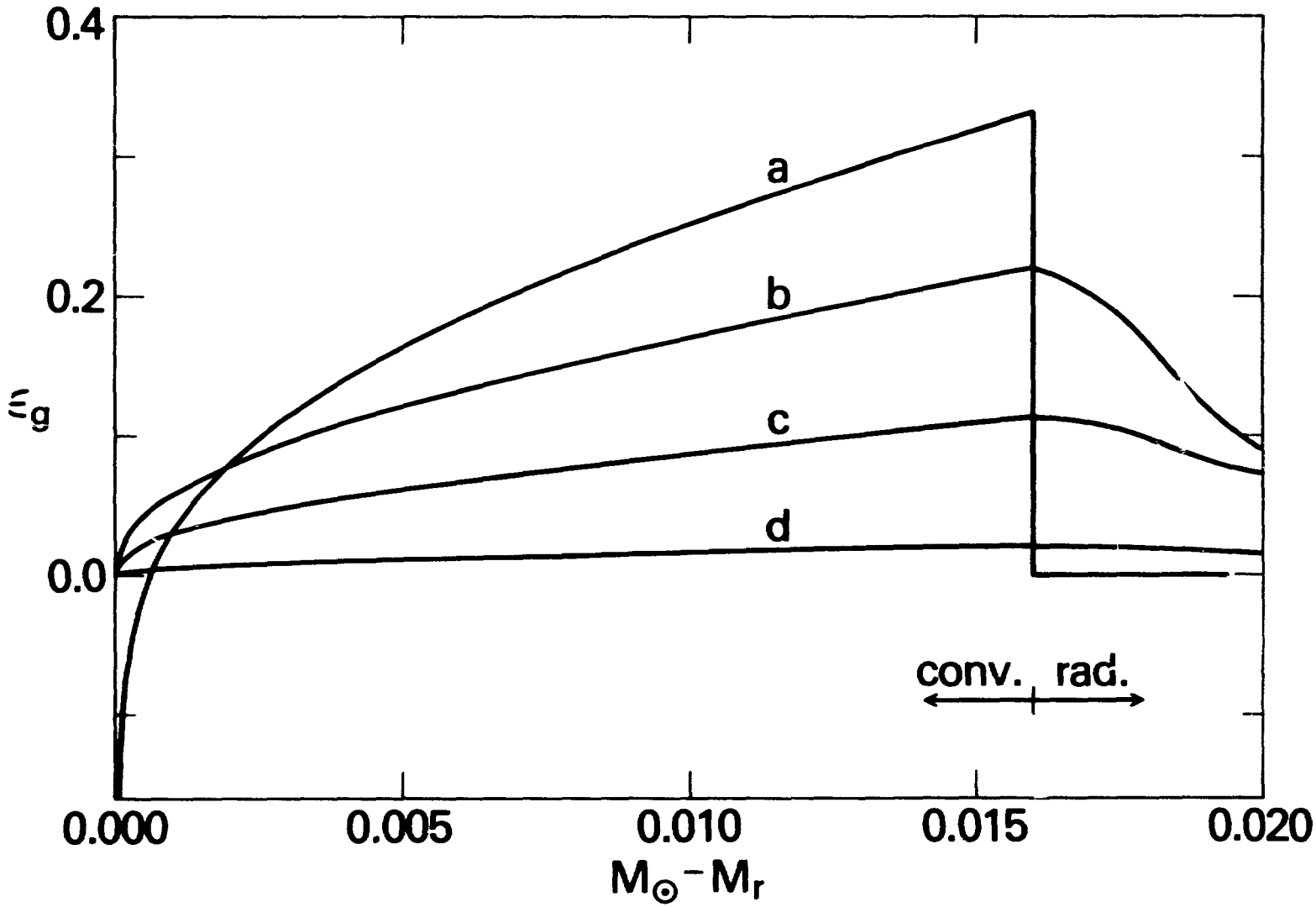


Figure 3. Rate of release of gravitational potential energy ε_g as a function of $M_\odot - M_r$ within the solar convective envelope at four times following the perturbation $\Delta\alpha = 0.01$

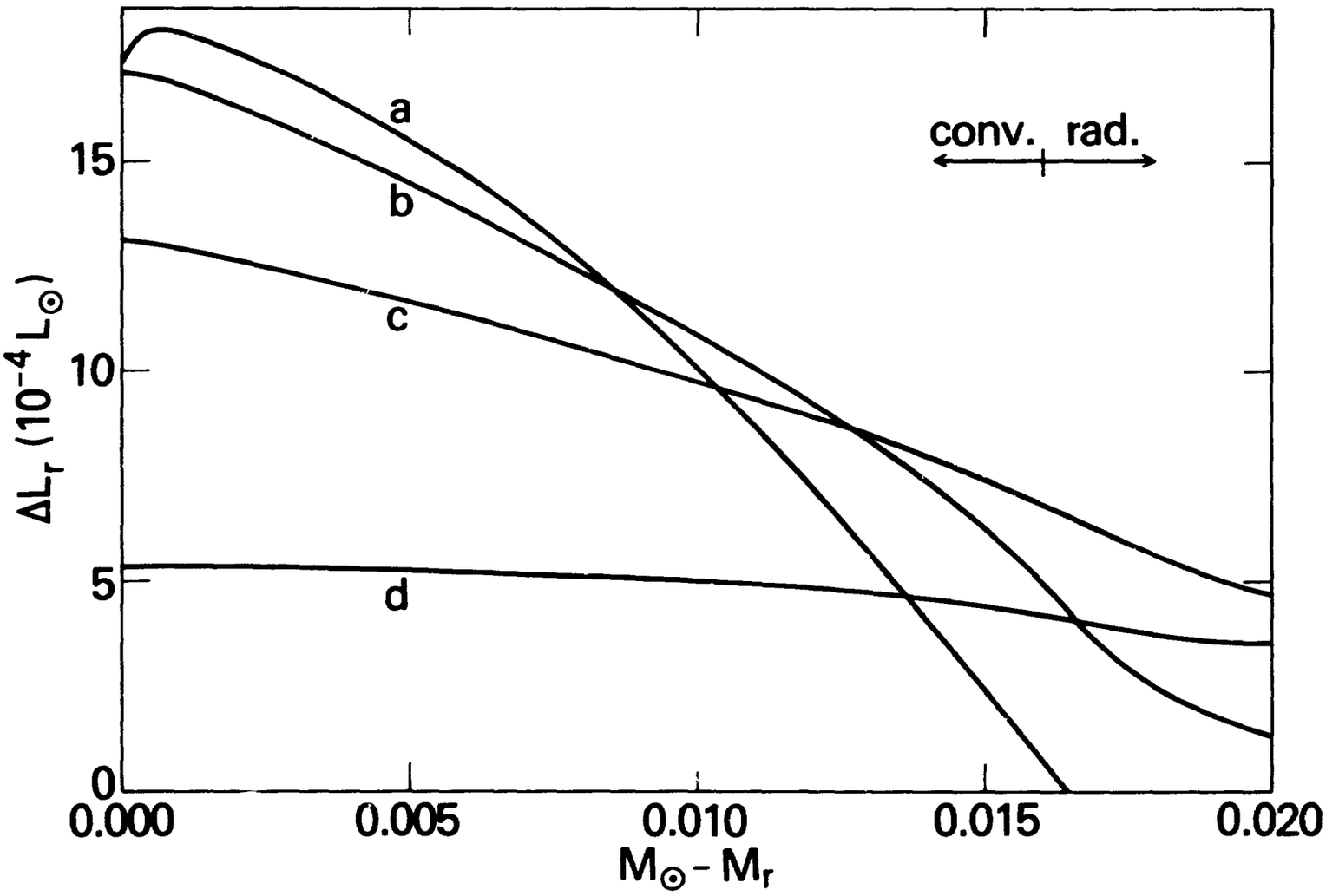


Figure 4. Perturbation in the outward flux ΔL_r as a function of $M_\odot - M_r$ within the solar convective envelope at four times following the perturbation $\Delta\alpha = 0.01$

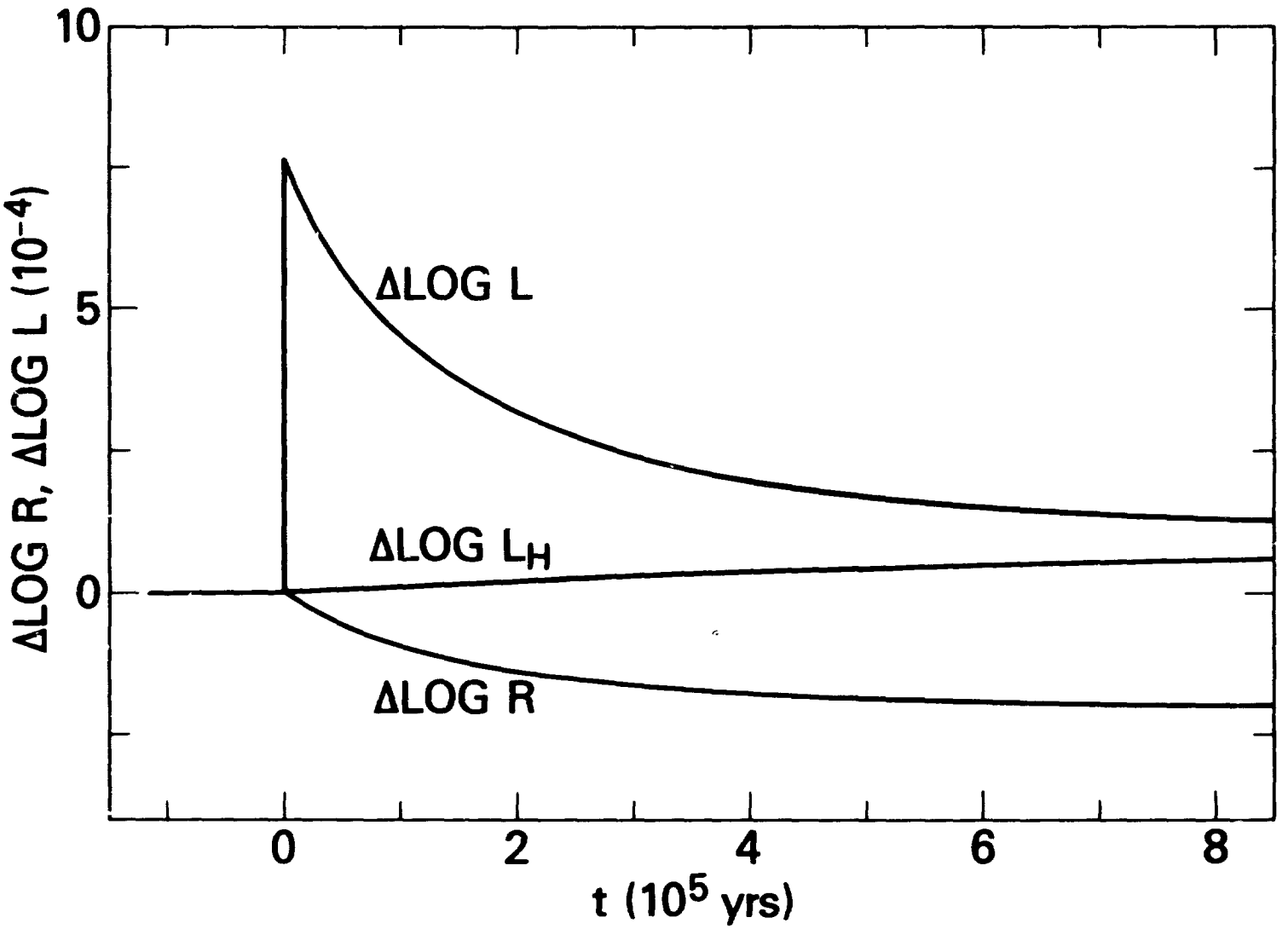


Figure 5. Time dependence of the change in the luminosity Δlog L, the surface radius Δlog R and the rate of hydrogen burning Δlog L_H following the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

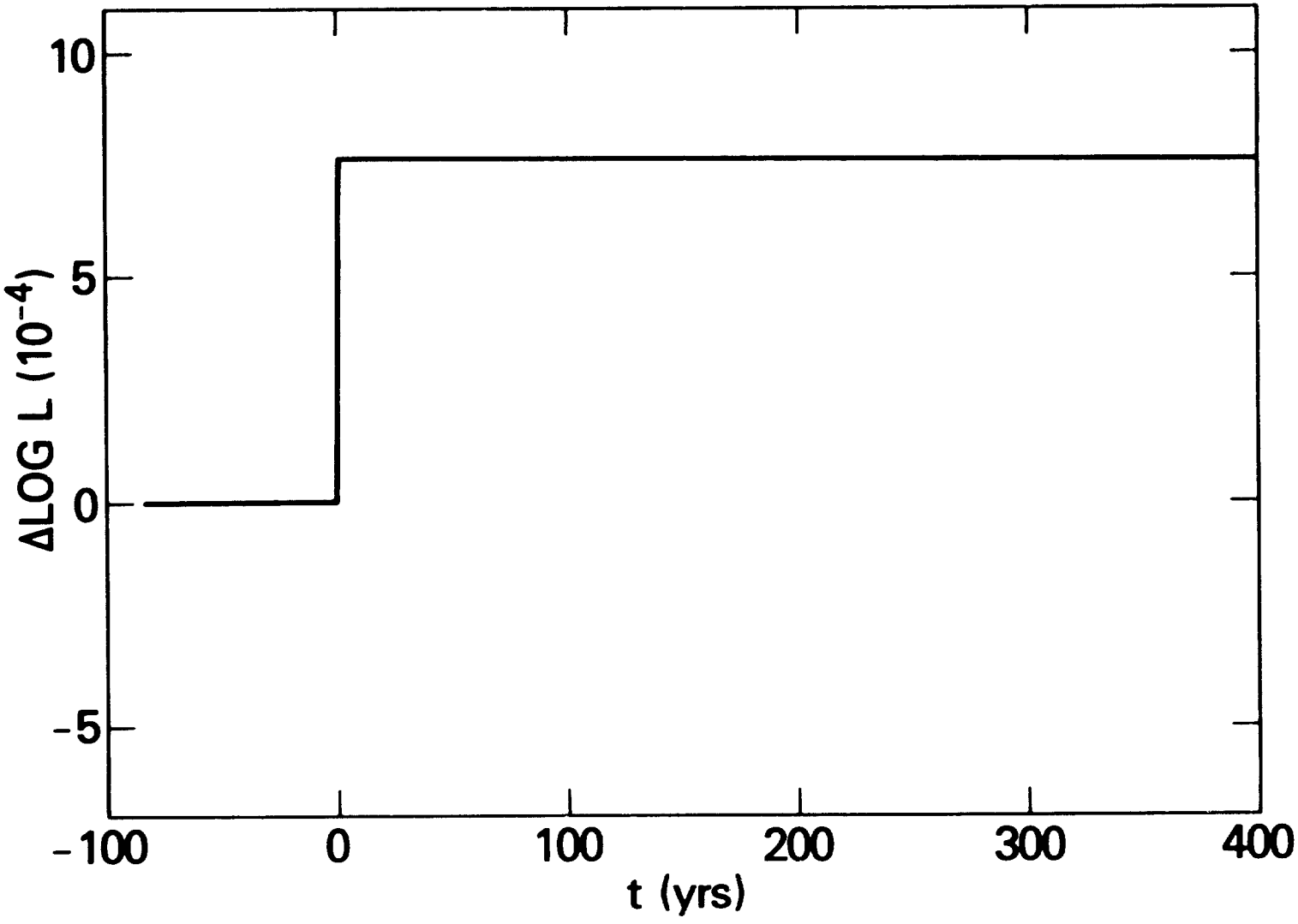


Figure 6. Time dependence of the change in the luminosity Δlog L shortly after the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

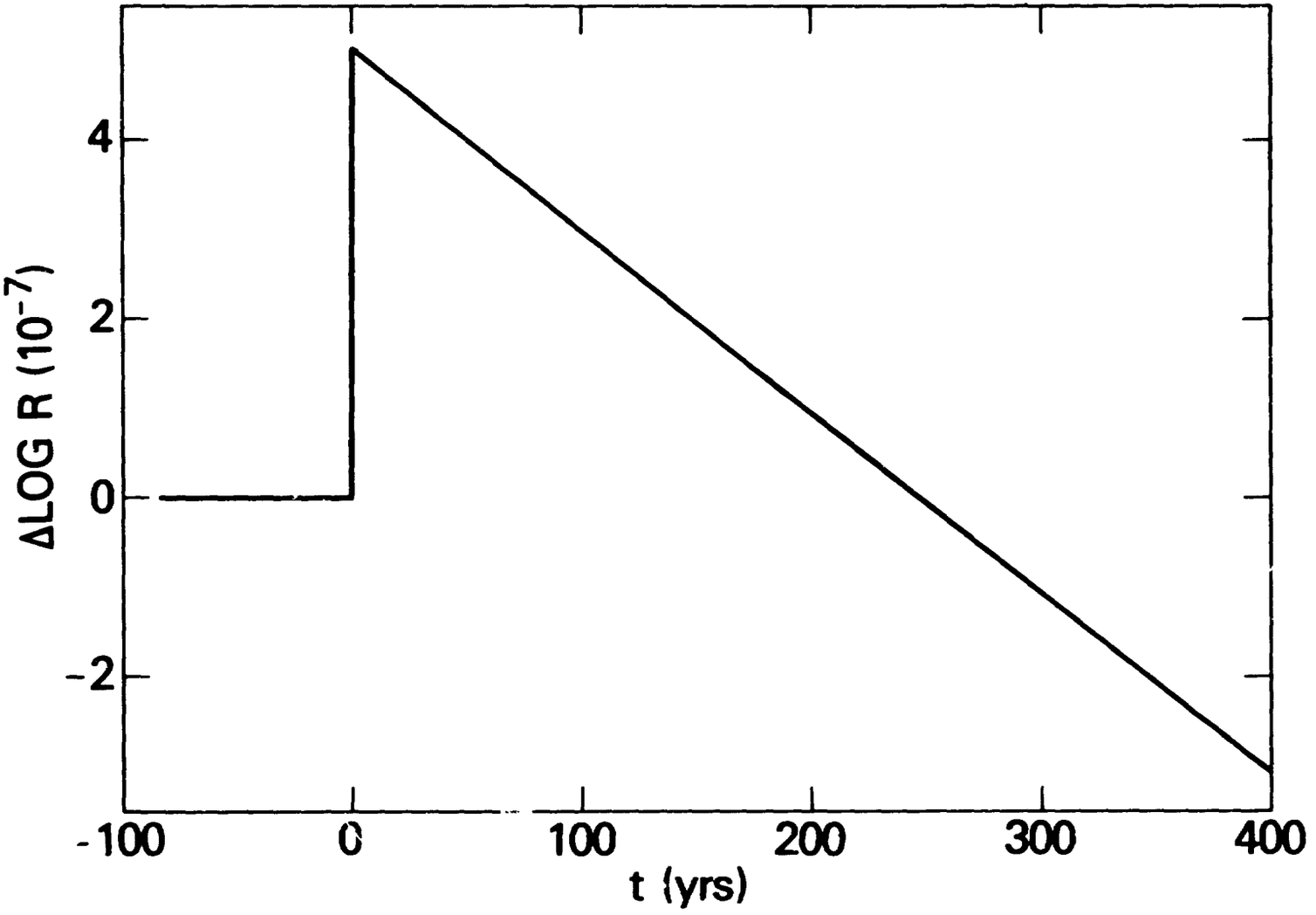


Figure 7. Time dependence of the change in the surface radius Δlog R shortly after the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

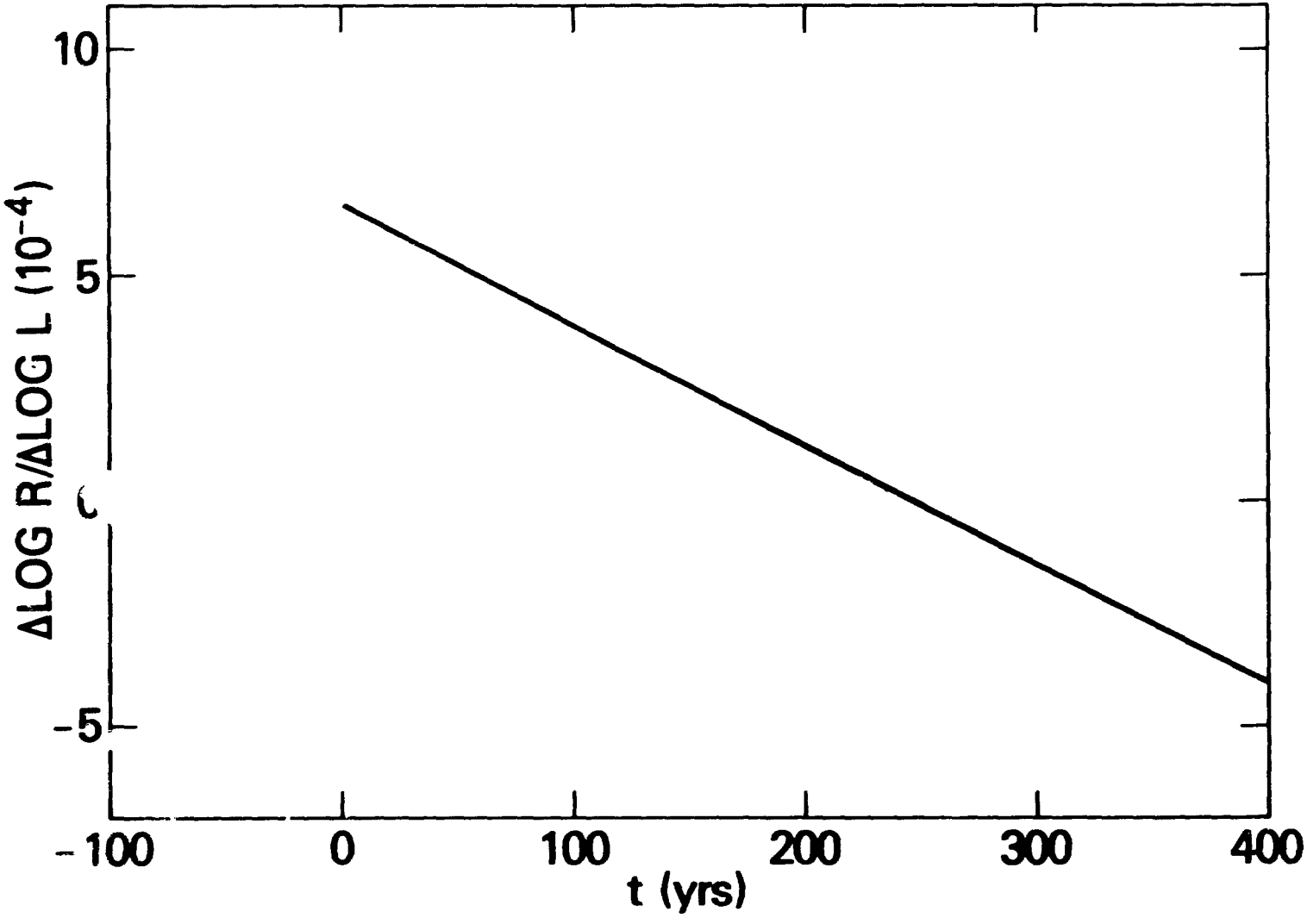


Figure 8. Time dependence of the ratio W = Δlog R/Δlog L shortly after the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

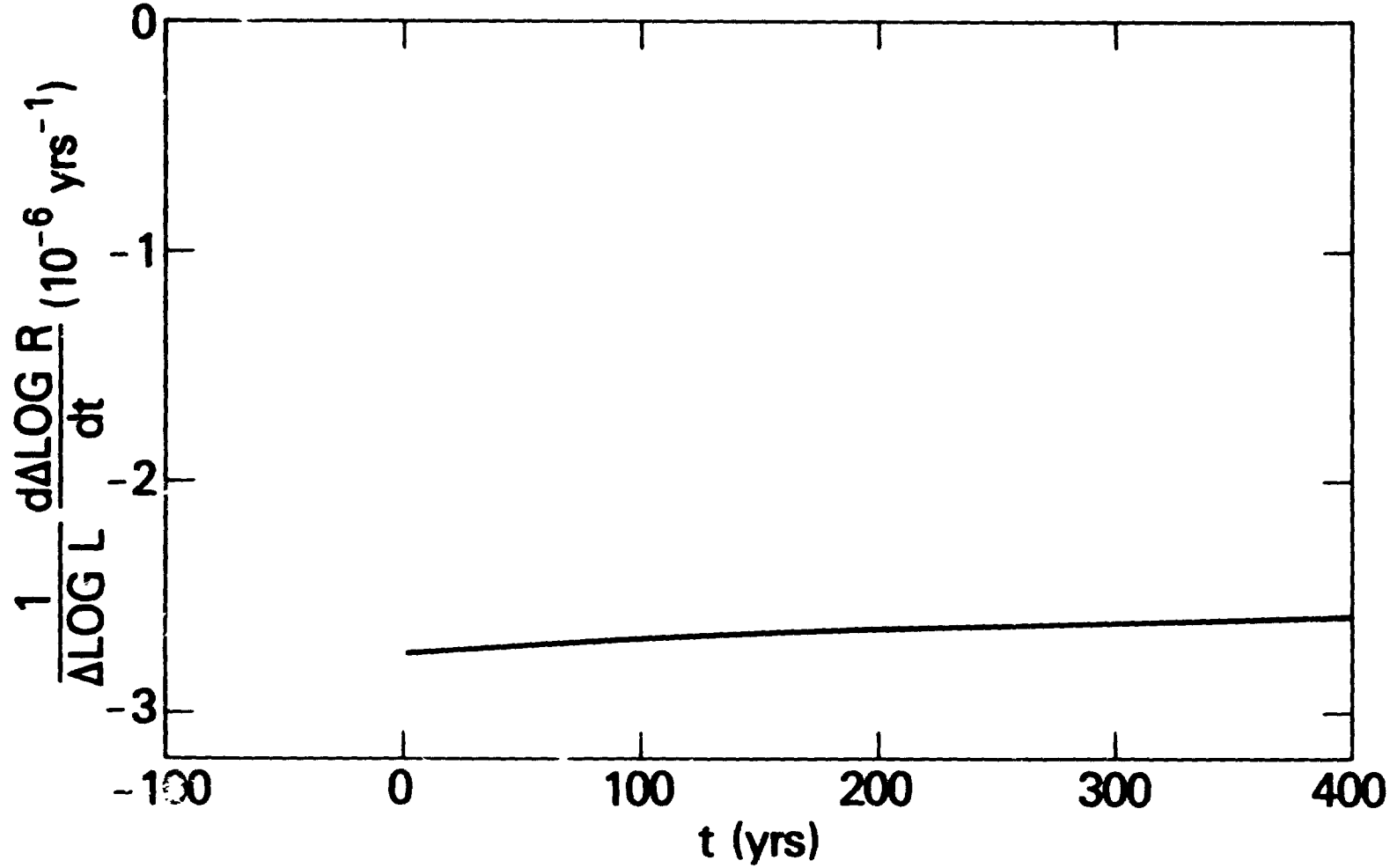


Figure 9. Time dependence of the ratio $H = (\Delta \log L)^{-1} d\ \Delta \log R/dt$ shortly after the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

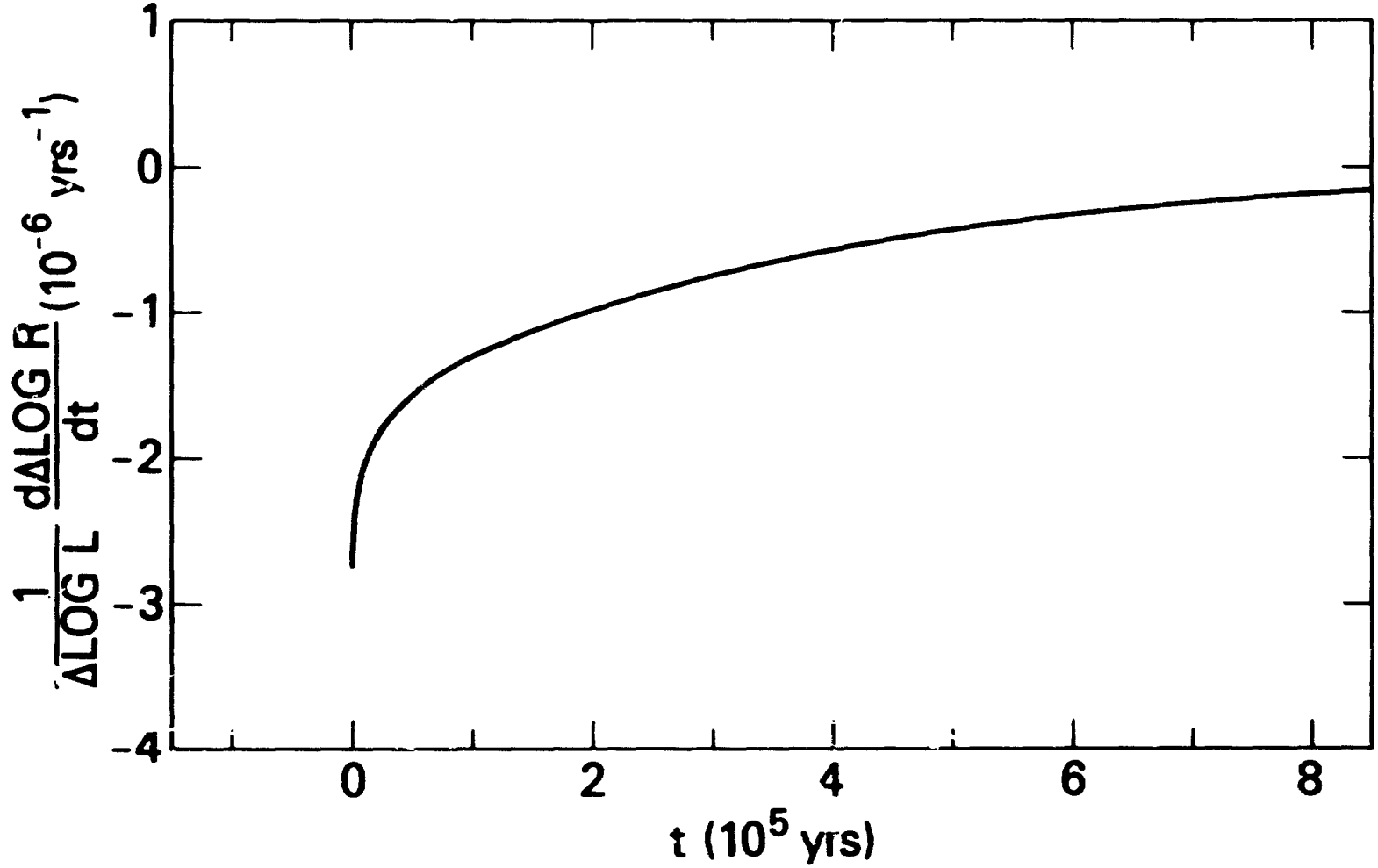


Figure 10. Time dependence of the ratio $H = (\Delta \log L)^{-1} d\ \Delta \log R/dt$ following the perturbation $\Delta\alpha = 0.01$ within the solar convective envelope

AN INSTRUMENT TO MEASURE THE SOLAR SPECTRUM FROM 170 TO 3200 NM ON BOARD SPACELAB*

G. Thuillier
Service d'Aéronomie du CNRS

P. C. Simon and R. Pastiels
Institut d'Aéronomie Spatiale de Belgique

D. Labs
Landessternwarte

H. Neckel
Hamburger Sternwarte

ABSTRACT

This instrument, at the present time in development, will fly on board Spacelab I in May 1983. Other flights are foreseen during the following missions.

The instrument is composed of three double monochromators covering the range 170 to 3200 nm. The spectrometers have bandpasses of 1 nm up to 900 nm and 20 nm from 850 to 3200 nm with an accuracy 10^{-2} nm. Calibration lamps are included in the instrument to monitor any change of its sensitivity and wavelength scale.

I. INTRODUCTION

The Earth has experienced several climatic changes. Those changes have been observed to have different amplitudes and time scales.

The causes of climatic changes are not as clearly identified as the different processes (thermodynamic, absorption of solar photon...) which act on a complex system composed of the solid Earth, Ocean, and Atmosphere.

It appears that for such a system, the Sun represents the most important external source of energy. Consequently, whatever the feedback processes existing in the Earth-atmosphere system are, a modification of the Sun input should lead to climate changes with a time delay that models could predict.

Some evidences of variation of the sun irradiance have been obtained, mainly at wavelength shorter than 200 nm and recently on the solar constant.

Then as several processes are wavelength dependent (photodissociation, ...), the scientific aims of the investigation are:

a. the measurement of the absolute solar spectral irradiance in the wavelength interval 180-3200 nm.

   The requested accuracy in general admitted for such measurement is 5% around 200 nm and 1% in the visible and IR ranges.

b. the measurement of possible long-term variation of the solar spectral irradiances.

   In this case, the requested precision is 0.3% in the visible and IR, and 1% in the UV.

Then, the proposed measurements are important in the three following fields:

1. Aeronomy of the middle atmosphere mainly concerned with the solar flux up to 400 nm,

2. Climatology concerned with the wavelength range above 400 nm necessary to establish the radiative budget of the atmosphere,

3. Solar Physics concerned with the full solar spectrum.

The requested accuracy and precision imply that the measurements are not achievable from the ground. Consequently, the measurement must be performed from space.

The vehicle chosen from this investigation is Spacelab launched by the Shuttle on account of:

- the mission is short and limits the time in which the instrument is exposed to space environment, which has always a degrading action,

- periodic missions being foreseen permit study of sun variation,

- the duration of the Shuttle program is compatible with a study conducted along a solar cycle,

- the return of the instrument permits a recalibration after flight which is quite necessary to reach the accuracy and precision needed by the investigation.

II. THE INSTRUMENT

It includes:

- 3 spectrometers
- 3 detectors

- Electronics
- 5 in-flight calibration lamps.

A schematic is given in Figure 1.

2.1 Spectrometers

The instrument includes three double monochromators of 10 cm focal length using concave holographic gratings. Their characteristics are:

- the six gratings are mounted on the same mechanical shaft rotating with a screw and nut system, giving a precision of 2 arc-second at any position. The spectral precision is 10^{-2} nm.
- the UV monochromator range of measurement overlaps with the one of the visible monochromator. Same as for visible and IR monochromators.
- transmitting diffusers (grinds) are placed in front of the three spectrometers.
- filter wheels are placed at the exit slits to remove the second order signal.
- wavelength range and spectral data:

Monochromators	Range	Band-pass	Lines/mm
UV	160 - 365 nm	1 nm	3600
VIS	277 - 889 nm	1 nm	1281
IR	805 - 3160 nm	20 nm	354

The principle of one spectrometer is given in Figure 2. A complete solar spectrum from 170 to 3200 nm is recorded in 15 minutes.

2.2 Detectors

UV: photomultiplier tube (EMR 641E)
VIS: photomultiplier tube (EMR 641E) cooled at +5°C
IR: PbS cell cooled at -10 C with a Chopper working at 512 Hz.

2.3 In-flight Calibration Lamps

The used detectors are not absolute detectors. Hence, the instrument has to be calibrated. But its sensitivity has to remain unchanged all along the ground operations before flight (which may be 1 or 2 years long for the first) and during the flight. Then, rather than maintain all the instrument stable in the above situations (which is difficult to achieve), it is easier to maintain stable a calibration lamp, the distance filament to slit and the lamp power supply. This philopophy has been adopted for this instrument.

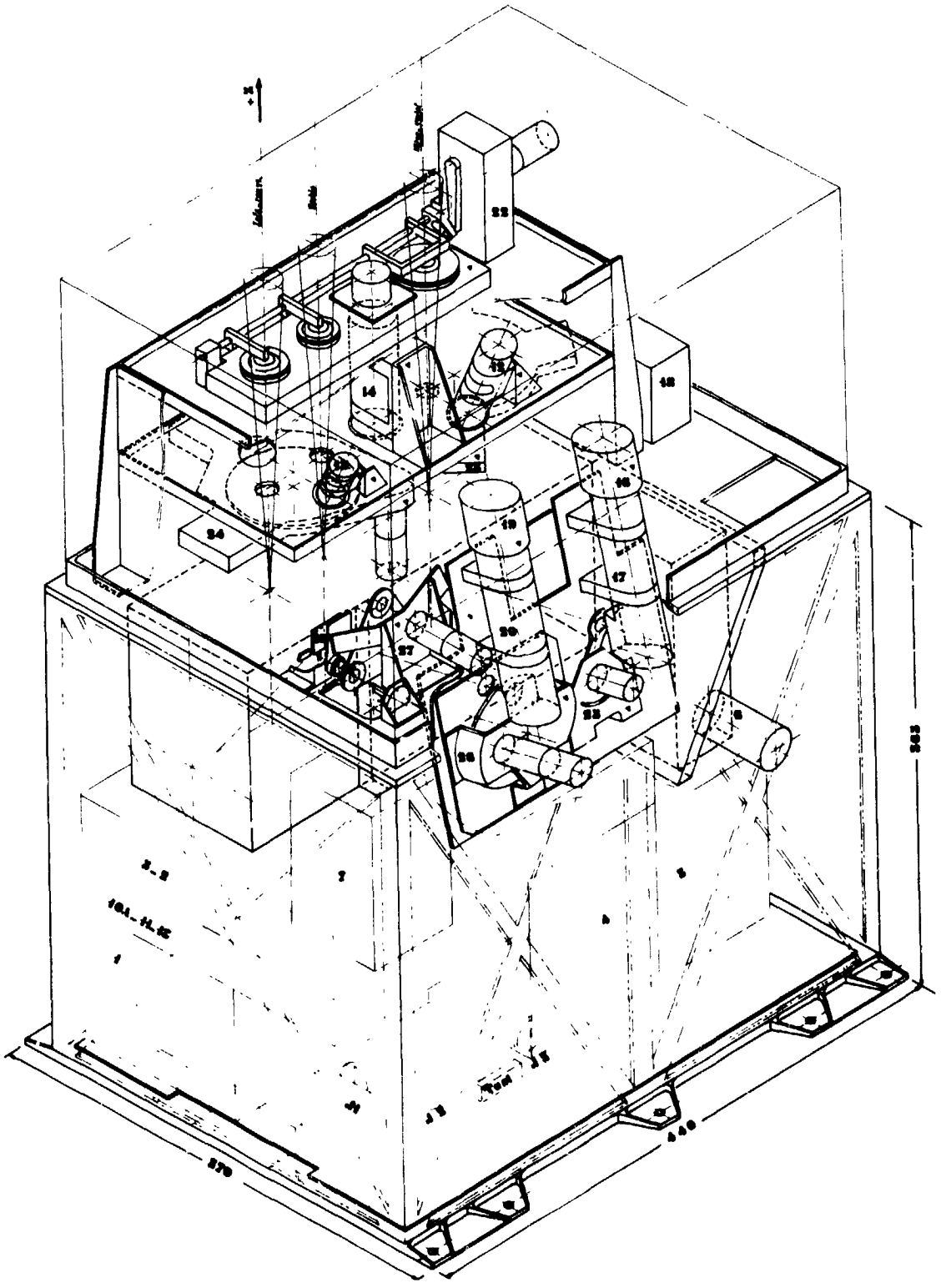


Figure 1. Schematic of the complete instrument

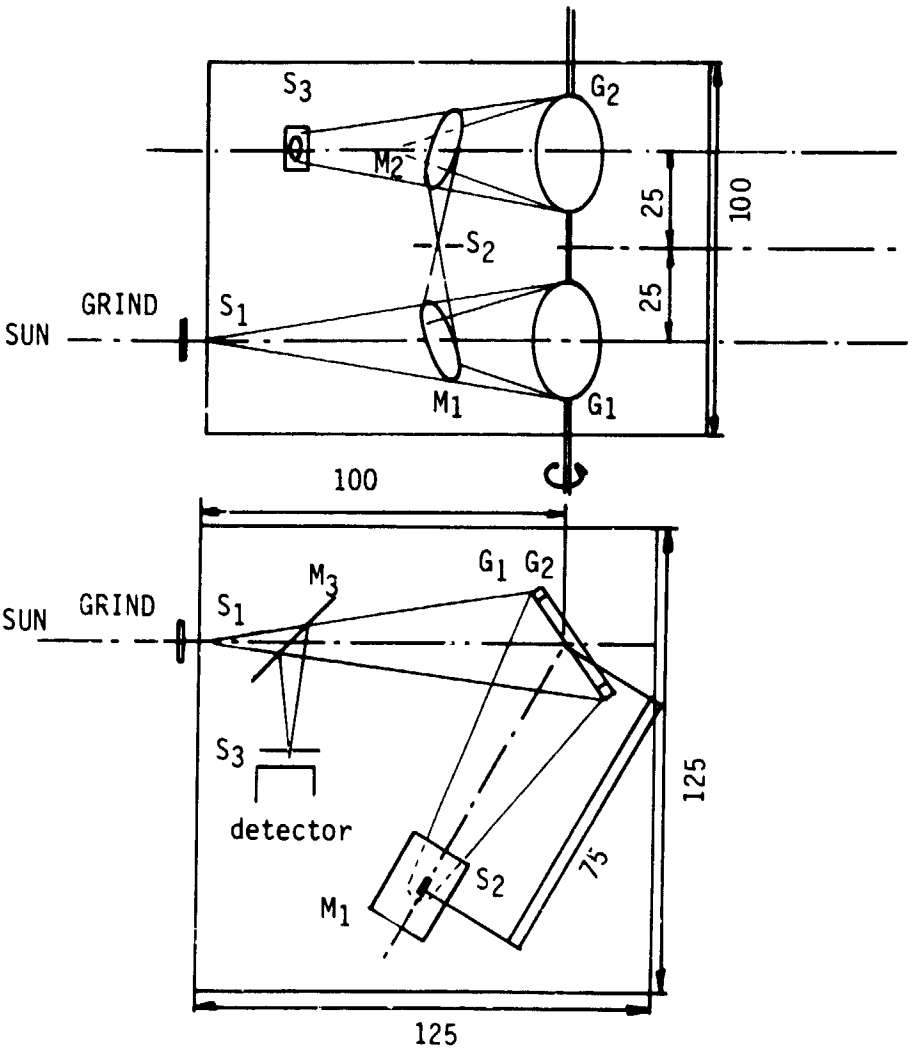


Figure 2. Schematic of one double monochromator. Three same units are mounted on the same mechanical shaft.

The following set of lamps is included in the instrument:

Lamp type	Spectrometer	Purpose	Power	Number
Tungsten ribbon	VIS and IR	Instrument sensitivity	5 W	2
Deuterium	UV		15 W	2
Hollow cathode (filled with He)	UV and VIS	Wavelength scale and band-pass	2 W	1

and has the general characteristics:

- the radiating region (ribbon and anode) is imaged by lenses on the input grinds limited by preslits,
- the diameter and magnification of the lenses are chosen so that the signal due to the lamps is of the order of the one due to the Sun,
- lamps and lenses are mechanically stable with respect to the entrance slits,
- the lamp currents are regulated at 10^{-4} to provide 0.1% stability,
- the lamp currents, voltage and temperature are telemetered,
- other characteristics are given in Ref. 1.

Consequently, we measure in space and at ground all possible change of the instrument sensitivity. The wavelength scale and bandpasses are also determined by use of selected lines emitted by the hollow cathode lamp. In this mode the rotation of the grating is performed by elementary step of 15 arc-second corresponding to 10^{-1} nm.

III. CALIBRATION OF THE INSTRUMENT

3.1 Absolute Calibration

The black body of the Landessternwarte Laboratory is used (3000 K). A linear pyrometer monitors the temperature. This device uses a set of interferential filters and a silicon cell which are temperature stabilized (0.1°C).

The accuracy of the calibration is estimated to be:

$\lambda < 250$ nm	5%
$250 < \lambda < 300$ nm	3%
$300 < \lambda < 400$ nm	2%
$\lambda < 400$ nm	1%

3.2 Wavelength Calibration

Its purpose is the measurement of the instrument bandpasses and the calibration of the wavelength scale.

Hollow cathode lamps filled with Argon or Neon are used. Other elements such as N_i, S_e, R_u, H_g, C_e, S_n, A_s, B_a, T_e, S_i, B_i, C_d permit covering the range from 170 to 3200 nm.

A tunable laser is also used.

3.3 Long Term Standard (LTS)

The scientific aim (b) needs a stable reference during some years, typically a half solar cycle. This requirement can be achieved by use of lamps since they have the needed stability.

The LTS consists of:

10 Tungsten ribbon lamps (W17G from Osram) for $\lambda > 250$ nm and 10 Deuterium lamps (D60 from Hanau) for $\lambda < 250$ nm. Five of each are carefully kept in laboratory while the remaining set are on special equipment named LTS box, able to follow the instrument at different sites for integration, preparation to launch and after mission. The LTS box can be placed on the top of the main instrument (Figure 3).

The precision of the LTS is:

1. 0.2% for 100 hours burning time with alternate DC power supply for tungsten lamps as it appears from measurements from 1960's in the Landessternwarte laboratory of Heidelberg.
2. 3% for 100 hours burning time with deuterium lamps (Ref. 2).

But, any change in their irradiance can be monitored by comparison with the tungsten ribbon lamp in the range 280-350 nm.

The set kept in laboratory will not be used more than 10 hours per year.

The sequencing of calibration and LTS use is given in Figure 4, where it appears that the LTS box will be calibrated against the black body using the space instrument itself.

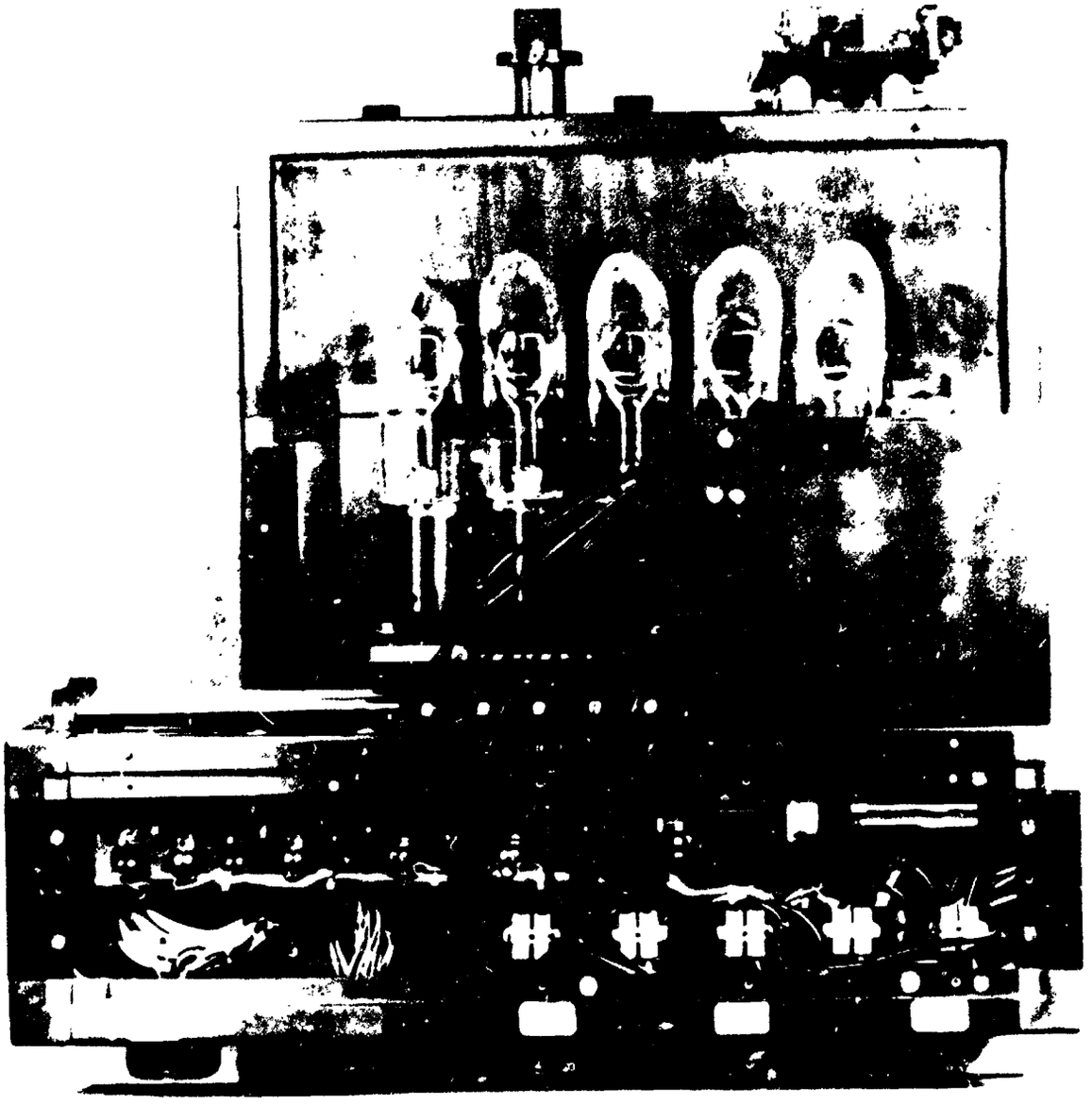


Figure 3. Long Term Standard box showing the five deuterium and five tungsten ribbon lamps. The lamps can be moved to light the three entrance slits of the monochromators.

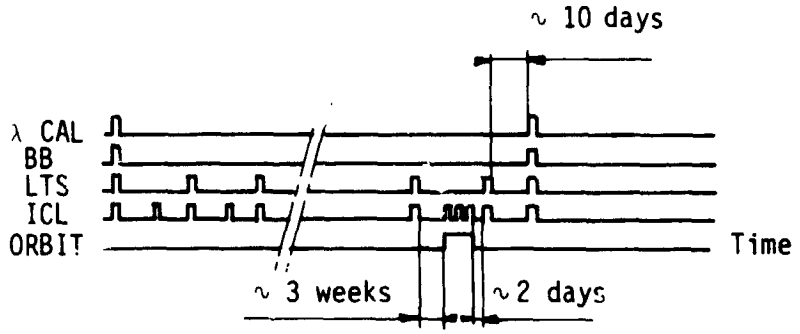


Figure 4. Sequence of calibration and use of LTS box
λ Cal: Bandpass and wavelength calibration
BB: Black body calibration
LTS: Long Term Standard
ICL: Use of In-flight Calibration Lamps.

3.4 Comparisons

Comparisons before flight with the two absolute radiometers also onboard Spacelab I is foreseen.

Other comparisons with calibrated sources such as the ones existing at the National Bureau of Standards are desired.

3.5 Accuracy and Precision

Accuracy has been given in Section 3.1.

The final precision of the instrument is estimated to be 0.3% taking into account the counting rate, the stability of the in-flight calibration and LTS lamps.

REFERENCES

1. Finkenzeller, U., and Labs, D.: The deuterium lamp as a UV continuum source from 180 nm - 320 nm for space application, Appl. Optics, 18, 3938, 1979.

2. Bridges, J. M., Ott, W. R., Pitz, E., Schulz, A., Einfeld, E., and Stuck, D.: Spectral radiance calibration between 165 and 300 nm; an interlaboratory comparison, Appl. Optics, 16, 1788, 1977.

D18

EVOLUTIONARY VARIATIONS OF SOLAR LUMINOSITY*

A. S. Endal
Department of Physics and Astronomy
Louisiana State University

ABSTRACT

Theoretical arguments for a 30% increase in the solar luminosity over the past 4.7 billion years are reviewed. A scaling argument shows that this increase can be predicted without detailed numerical calculations. The magnitude of the increase is independent of nuclear reaction rates, as long as conversion of hydrogen to helium provides the basic energy source of the Sun.

The effect of the solar luminosity increase on the terrestrial climate is briefly considered. It appears unlikely that an enhanced greenhouse effect, due to reduced gases (NH_3, CH_4), can account for the long-term paleoclimatic trends.

INTRODUCTION

Climatically significant changes of the solar luminosity (L) have been postulated to occur on time scales ranging from a few years to billions of years. The shorter time scales have been discussed extensively at this conference. In the present review, I will restrict myself to the longest time scales ($\gtrsim 10^9$ yr.) and discuss the basis for the astrophysical conclusion that the Sun was $\sim$ 30% fainter 4.7×10^9 yr. ago and that the evolution since the Sun's formation requires a slow, but steady, increase in L.

I should note that this is the only change in L predicted by stellar evolution theory, in its standard form. This prediction is common to all modern calculations and is supported by a large body of data from observational stellar astronomy (see reference 1 for a review of the observational evidence). Nevertheless, the validity of this result has been questioned because of the apparent conflict with proxy indicators of the Earth's past climate (ref. 2-4). For this reason, a review of the theoretical arguments for the long-term increase of L is in order.

Stellar evolution is governed by nonlinear differential equations derived from conservation laws and considerations of energy transport processes. Analytic solutions do not exist for any cases relevant to the Sun so numerical solutions must be used. Modern calculations require complex computer codes incorporating a variety of physical data on nuclear parameters, transport coefficients, and thermodynamic properties. In this respect, the situation is

*Research supported in part by NASA grant NAG 5-13 and NSF grant AST 79-19688.

similar to that encountered in current theoretical investigations of the terrestrial climate. This may, in fact, explain the reluctance of the climatologists to accept the astrophysical result: climatologists understand the pitfalls of accepting solutions obtained from complex computer codes at face value. For this reason, I will largely avoid discussion of numerical models and base the astrophysical case on simple scaling laws.

A SCALING MODEL OF THE SUN

We begin by requiring that the Sun be in hydrostatic equilibrium, with gravitational forces balanced by the pressure gradient. The free-fall time of the Sun is on the order of an hour and any departures from hydrostatic equilibrium would show up as luminosity and radius changes on this time scale. For the spherically symmetric case, hydrostatic equilibrium is expressed as

$$\frac{dP}{dr} = -\frac{Gm_r}{r^2}\rho , \tag{1}$$

where G is the gravitational constant, P and ρ are the pressure and density at a distance r from the center and m_r is the mass interior to r. Measurements of the visible solar disk show that the Sun is spherical to within 1 part in 10^5 (ref. 5).

We can construct a one-zone model by replacing (1) by a finite-difference equation evaluated between the center c and surface s, with mean values enclosed in brackets $\langle\ \rangle$:

$$\frac{P_c - P_s}{r_c - r_s} = -G\left\langle \frac{m_r\rho}{r^2} \right\rangle .$$

Applying the boundary conditions $P_s = 0$, $r_s = R$ (radius), and $r_c = 0$ gives

$$P_c = G\left\langle \frac{m_r\rho}{r^2} \right\rangle R. \tag{2}$$

The scaling laws for the mean values are:

$$\langle m_r \rangle \propto m, \tag{3}$$

$$\langle r \rangle \propto R, \tag{4}$$

and

$$\langle \rho \rangle = \frac{m}{\frac{4}{3}\pi R^3} \propto \frac{m}{R^3} , \qquad (5)$$

where m is the total mass. Inserting these scaling laws into equation (2) gives

$$P_c \propto \frac{m^2}{R^4} . \qquad (6)$$

To proceed further, we need an equation of state, relating the pressure to the density and temperature (T). For typical conditions characterizing the bulk of the solar interior ($\rho \simeq 1$ g/cm^3 and $T \simeq 10^6$ to 10^7 ^oK), Coulomb interaction energies are at least 2 orders-of-magnitude smaller than particle kinetic energies. Thus, the ideal gas law is an excellent approximation and this is what differentiates a star from a planet. Applying the ideal gas law to the center gives

$$P_c = \frac{N_A k}{\mu} \rho_c T_c , \qquad (7)$$

where k is the Boltzmann constant and the particle density is expressed as the Avogadro number N_A divided by the mean mass μ (in atomic mass units) per free particle. Eliminating P_c between (6) and (7) and noting that the central density must scale as the mean density gives

$$T_c \propto \frac{\mu m^2}{R^4} \left\langle \frac{1}{\rho} \right\rangle$$

If the scaling law (5) is used to replace R, we get

$$T_c \propto \mu m^{2/3} \langle \rho \rangle^{1/3} . \qquad (8)$$

We now turn to the question of how energy is transported from the core, where nuclear reactions produce energy, to the surface. Due to the high temperatures, radiative transport of energy is very efficient and dominates over the bulk of the interior. The mean free path ℓ of a photon is typically 1 cm so the photon diffusion approximation is valid to order $\ell/R \simeq 10^{-11}$. The radiative diffusion equation with spherical symmetry is

$$l_r = \frac{64\pi\sigma}{3} \frac{r^2 T^3}{\kappa\rho} \frac{dT}{dr} , \qquad (9)$$

where L_r is the total flux across a spherical surface at distance r from the center, σ is the Stefan-Boltzmann constant and κ is the Rosseland-mean opacity coefficient. Again, we use the one-zone difference approximation to write this equation as

$$\langle L_r \rangle = \frac{16\pi\sigma}{3} \left\langle \frac{r^2 T^3}{\kappa\rho} \right\rangle \frac{T_c - T_s}{r_c - r_s} \quad .$$

Application of our previous boundary conditions plus $T_s = 0$ (i.e. $T_s << T_c$) gives

$$L \propto \frac{R^2 T^3}{\kappa\rho} \frac{T}{R} = \frac{RT^4}{\kappa\rho} \tag{10}$$

where, since we are now dealing with a scaling law (proportionality), L_r can be replaced by L, r by R, etc.

To evaluate the opacity coefficient κ, we note that, in the solar interior, hydrogen and helium will be completely ionized and the heavier ions will be stripped of most of their electrons. Hydrogen and helium affect κ through free-free transitions while the heavier elements contribute primarily through bound-free transitions. Both processes are reasonably represented by the hydrogenic approximation so the absorption coefficient for a given ion varies inversely with the cube of the frequency. Although individual ionization states may contribute "noise" to the detailed dependence of κ on ρ and T, the broad dependence is given by Kramers' opacity:

$$\kappa = \kappa_o \rho\, T^{-3.5} \quad . \tag{11}$$

Putting this result into equation (10), and using equation (8) to eliminate the temperature, gives

$$L \propto m^{5.33}\, \rho^{0.17}\, \mu^{7.5} \quad . \tag{12}$$

The present rate of mass loss, due to the solar wind, is roughly $10^{-14} m_o$/yr. (ref. 6) and there is no reason to believe that the mass loss rate in the past was great enough to significantly affect m. The density dependence in (12) is so weak that we may also neglect changes in this parameter. Thus, the luminosity is primarily dependent on the mean molecular weight μ and we rewrite (12) as

$$L(t) = L(o) \left[\frac{\mu(t)}{\mu(o)}\right]^{7.5} \tag{13}$$

C - 3

We let X, Y, and Z denote the fractional abundances, by mass, of hydrogen, helium, and metals (X+Y+Z = 1). For a fully ionized gas,

$$\mu \simeq \frac{1}{2X + \frac{3}{4}Y + \frac{1}{2}Z} = \frac{2}{1 + 3X + \frac{1}{2}Y} \qquad (14)$$

The mean molecular weight increases with time due to conversion of hydrogen (μ = 1/2) into helium (μ = 4/3), by nuclear reactions, producing Q = 6×10^{18} erg per gram of hydrogen consumed. This energy must supply the luminosity of the Sun. Since Xm is the total mass of hydrogen,

$$\frac{d(Xm)}{dt} = m \frac{dX}{dt} = -\frac{L}{Q} ,$$

or

$$\frac{dX}{dt} = -\frac{L}{mQ} \qquad (15)$$

Differentiating (14) with respect to time, and noting that dY/dt = - dX/dt, we get

$$\frac{d\mu}{dt} = -\frac{5}{4}\mu^2 \frac{dX}{dt} = \frac{5}{4}\frac{\mu^2 L}{mQ} . \qquad (16)$$

Finally, we can eliminate μ between equations (13) and (16). The resulting differential equation can be directly integrated to give

$$L(t) = L(o)\left[1 - \frac{85}{8}\frac{\mu(o)\, L(o)}{mQ}\, t\right]^{-15/17} \qquad (17)$$

Since nuclear reactions are confined to the core, the present photospheric abundances should reflect the initial composition. Thus, we may evaluate μ_o using X $\simeq$ 0.71 and Z = 0.02. Equation (17) becomes

$$L\left(\frac{t}{t_\odot}\right) = L(o)\left[1 - 0.35\, L(o)\left(\frac{t}{t_\odot}\right)\right]^{-15/17} \qquad (18)$$

where L is expressed in units of the present solar luminosity (taken as 3.9×10^{33} erg/s) and $t_\odot$ is the present solar age (4.7×10^9 yr.). The initial luminosity required to match the present solar luminosity at $t_\odot$ is L(o) = 0.76.

The scaling arguments predict that the Sun was initially 24% fainter than the present luminosity. A comparison of the luminosity evolution according to equation (18) with results from detailed numerical models (ref. 7) is shown in figure 1. As noted by D. O. Gough, the evolution predicted by numerical models is accurately represented by

$$L = \left[1 + \frac{2}{5}\left(1 - \frac{t}{t_\odot}\right)\right]^{-1} L_\odot , \qquad (19)$$

where $L_\odot$ is the present solar luminosity. This formula, rather than equation (18), is recommended for studies of the evolution of planetary atmospheres.

SUMMARY OF THE ASTROPHYSICAL CASE

The above analysis shows that a quantitative prediction of the evolutionary increase of the Sun's luminosity may be made without detailed knowledge of the physical processes taking place in the interior. Therefore, this prediction is not affected by the uncertainties in this knowledge. In particular, we did not have to specify any nuclear reaction rates since the net reaction rate, integrated over the solar mass, is determined by the measured solar luminosity. This is quite different from the case of the solar neutrino prediction, which is very sensitive to detailed nuclear reaction rates (ref. 8). The discrepancy between the predicted and observed neutrino flux should not be used to argue that the luminosity prediction is also questionable.

APPLICATION TO THE EARTH'S CLIMATE

Sagan and Mullen (ref. 9) pointed out that an enhanced greenhouse effect, due to higher concentrations of NH_3 and CH_4 in the Earth's atmosphere, could have maintained a warm climate even with a lower solar luminosity. A similar conclusion was reached by Hart (ref. 10). This mechanism cannot, however, compensate for all of the solar luminosity evolution.

Paleological evidence (ref. 11) shows that the Earth's atmospheric chemistry changed from reducing to oxidizing some 1.5 to 2 billion years ago and this would have removed the enhanced greenhouse effect due to reduced compounds. Roughly one-half of the solar luminosity increase occurs during the last 2 billion years but there is no evidence for a parallel increase in the Earth's mean surface temperature. Indeed, isotopic studies of Precambrian samples by Knauth and Epstein (ref. 12) indicate that the mean surface temperature has been decreasing during this time. Clearly, there is a need for further studies of the effects of crustal movements and volcanism,

biological activity, etc. on the long-term evolution of the Earth's climate. At present, it appears that the effects of solar evolution are still buried in the "noise" due to other uncertainties in paleoclimatic models.

REFERENCES

1. Newkirk, G., Jr.: Solar Variability on Time Scales of 10^5 to $10^{9.6}$ Years. The Ancient Sun (ed. R. O. Pepin, J. A. Eddy, and R. B. Merrill), Geochim. Cosmochim. Acta, suppl. no. 13, 1980, pp. 293-320.

2. Pollack, J. B.: Climatic Change on the Terrestrial Planets. Icarus, vol. 37, no. 3, March 1979, pp. 479-553.

3. Toon, O. B., Pollack, J. B., and Rages, K.: A Brief Review of the Evidence for Solar Variability on the Planets. The Ancient Sun (ed. R. O. Pepin, J. A. Eddy, and R. B. Merrill), Geochim. Cosmochim. Acta, suppl. no. 13, 1980, pp. 523-531.

4. North, G.: Impact of Solar Constant Variations on Climate. Workshop on Solar Constant Variations, NASA CP-2191, 1981.

5. Hill, H. A., and Stebbins, R. T. : The Intrinsic Visual Oblateness of the Sun, Astrophys. J., vol. 200, no. 2, Sept. 1975, pp. 471-483.

6. Brant, J. C.: The Solar Wind. W. H. Freeman and Co., 1970.

7. Endal, A. S., and Sofia, S.: Rotation in Solar-Type Stars. Astrophys. J., in press, 1981.

8. Bahcall, J. N.: Solar Neutrinos: Theory Versus Observations. Space Sci. Rev., vol. 24, no. 2, 1979, pp. 227-251.

9. Sagan, C., and Mullen, G.: Earth and Mars: Evolution of Atmospheres and Surface Temperatures. Science, vol. 177, no. 4043, July 1972, pp. 52-56.

10. Hart, M. H.: The Evolution of the Atmosphere of the Earth. Icarus, vol. 33, no. 1, Jan 1978, pp. 23-39.

11. Cloud, P.: Beginnings of Biospheric Evolution and their Biogeochemical Consequences, Paleobiology, vol. 2, no. 4, Fall 1976, pp. 351-387.

12. Knauth, L. P., and Epstein, S.: Hydrogen and Oxygen Isotope Ratios in Nodular and Bedded Cherts. Geochim. Cosmochim. Acta, vol. 40, no. 9, Sept. 1976, pp. 1095-1108.

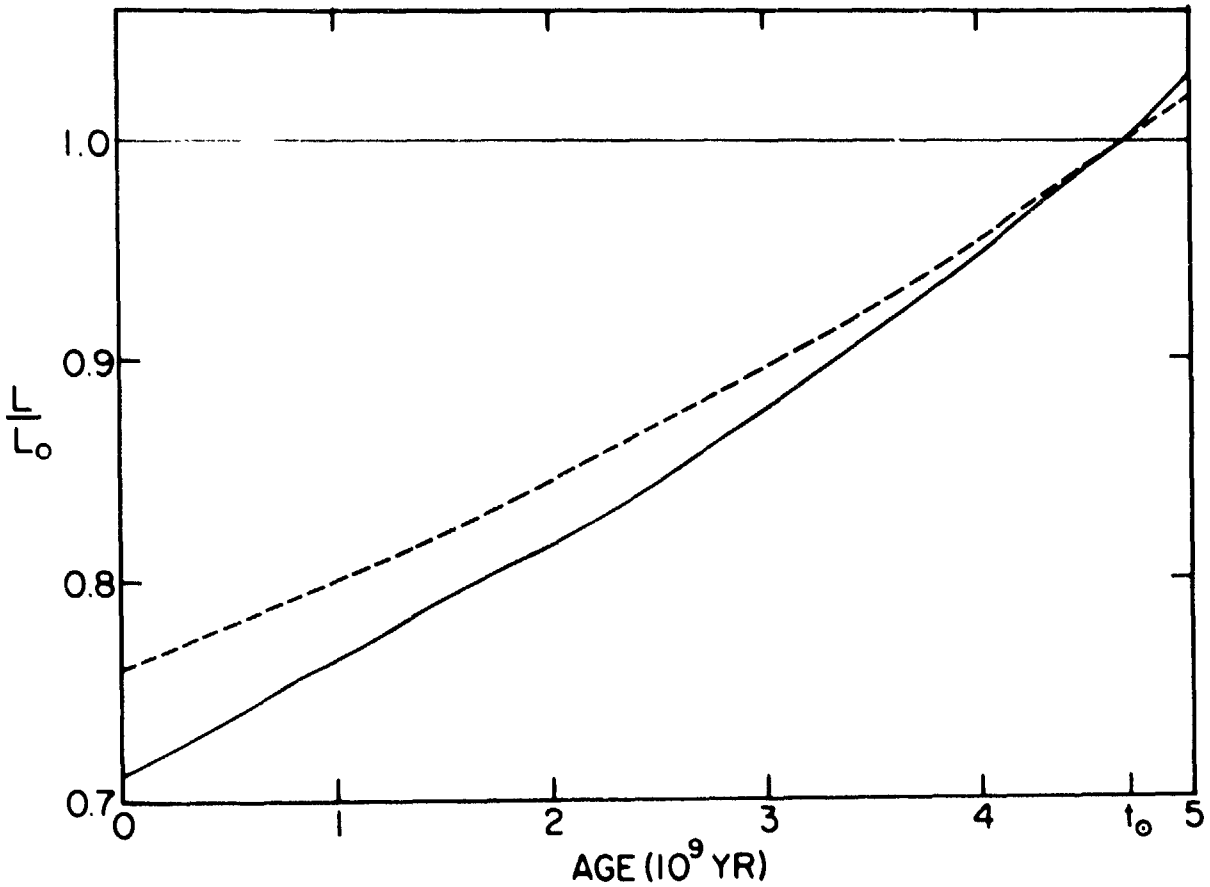


Figure 1. Long-term evolution of the solar evolution. The evolution predicted by the scaling model (Equation (18)) is shown by the dashed line and the prediction from a detailed computer model is shown by the solid line.

D19

ON THE SEAT OF THE SOLAR CYCLE

Douglas Gough
Institute of Astronomy, and Department of Applied Mathematics
and Theoretical Physics, University of Cambridge

ABSTRACT

This paper is a discussion of some of the issues that have been raised in connection with the seat of the solar cycle. Is the cycle controlled by a strictly periodic oscillator that operates in the core, or is it a turbulent dynamo confined to the convection zone and possibly a thin boundary layer beneath it? Sunspot statistics are discussed, with a view to ascertaining the length of the memory of the cycle, without drawing a definitive conclusion. Also discussed are some of the processes that might bring about variations δL and δR in the luminosity and the radius of the photosphere. It appears that the ratio $W = \delta \ln R / \delta \ln L$ increases with the depth of the disturbance that produces the variations, so that imminent observations might determine whether or not the principal dynamical processes are confined to only the outer layers of the sun.

INTRODUCTION

Early theories of the solar cycle were based on the idea that there is some periodic oscillation in which the entire sun participates. However, most solar physicists today probably believe that the cycle is the product of a turbulent dynamo in the convection zone. This belief appears to have been based originally on the premise that the solar interior could not possibly turn over in a time as short as 11 or 22 years, which would be necessary if the magnetic field external to the sun matched smoothly to the field beneath the convection zone. And the belief has been strengthened by the comparatively complicated theoretical edifice that has been erected to explain some of the observations in terms of a turbulent dynamo. That edifice has been of great use in helping us to understand the kinds of processes that are no doubt operative in the sun's convection zone, but one must be wary of taking too seriously the results of what are physically quite naive models. These models all neglect, often without serious discussion, what might be quite important phenomena, and one of those that may be of considerable interest is the coupling to the solar interior. It is to this issue that I intend to devote my discussion.

Aside from wanting to understand the solar cycle per se, a knowledge of the most important dynamical aspects is essential for any discussion of how, or whether, the cycle has any relevance to other issues that concern the sun. Does the mere existence of the cycle tell us anything about the conditions in the solar core? This question has been raised several times in connection with the solar neutrino problem, for example, and Dicke (1) has discussed it in connection with the $12^{d}.2$ modulation of the Princeton oblateness data.

I might say in passing that an obvious point of interest in the solar cycle is its relation to the earth, and its influence on the climate and the ^{14}C production. These are discussed in other contributions to these proceedings. The only point I wish to add is that if any firm relation between the solar behaviour and measurable terrestrial records can be established, then the records might give us a measure of that behaviour that extends further back in the past than direct solar observations. This would be of obvious importance for improving our knowledge of the statistics of the cycle, to which I now turn my attention.

STATISTICS OF THE SOLAR CYCLE

Though theories of the solar cycle that depend on oscillations of the entire sun have not reached the level of sophistication attained by dynamo theories, and therefore may seem at first sight less plausible because they immediately raise unanswered questions in the minds of anyone who considers them, it does not necessarily follow that the ideas behind them are incorrect. To some extent dynamo theory may have suffered* the Gold effect (2), and to rescue it from this plight one should stand back and ask just what the predictions of the competing hypotheses are, and whether one really can discriminate between them by comparison with observation. There are many discussions of this issue, including the excellent critique by Cowling (3). Here I simply take up a point that Dicke (4) has raised, and ask whether the observed departure of the cycle from a regular oscillation can be used as a test.

If the dynamics of the cycle is controlled by a perfectly regular oscillation of the solar interior, then the manifestation of that oscillation by the sunspots ought to be closely linked to the state of the interior. I am not concerned here with whether the interior oscillation is itself directly responsible for the generation of the magnetic field, or whether it merely controls the dynamo in the convection zone. All I ask is whether the epochs of sunspot maxima and minima are closely linked in phase with a perfect clock.

A modern example of a model with an almost perfect clock is that proposed by Dicke (1): magnetic field of alternating polarity is released periodically from the core, and then rises slowly to the surface to produce the sunspots. The rise time is variable (5, 6): on the whole it is shorter the greater the total flux, which is what one might expect from magnetic buoyancy arguments, and provides a natural explanation for the correlation between the early onset of a new cycle and the sunspot number at the next sunspot maximum; in addition there are random fluctuations in the rise time induced by the turbulence in the convection zone. Associated with the release of the field is a contemporary variation in luminosity, which is presumed to be strictly periodic, and which is proposed to be responsible for climatic variation. Thus it is the interior oscillation itself that should be observable in climatic records, and not the sunspots. What must be somewhat disturbing to any proponent of

---

* or, perhaps, enjoyed

this theory therefore, is Murray Mitchell's report at this conference that the mean US drought record correlates better with the solar magnetic cycle than it does with a strictly periodic oscillator.

By comparing the sunspot data with mean [D]/[H] ratios obtained from two bristlecone pines, averaged over samples representing non-overlapping ten-year growth intervals throughout a period of 1000 years, Dicke concluded that the mean rise time of the magnetic field is about 13 y. His argument is that the climatic record at the location of the pines, when viewed as ten-year averages, shows strong signs of a 22 y oscillation which might maintain phase, and whose maximum amplitude occurs about 2 + iP years prior to sunspot maximum, where P ≈ 11 y is the mean duration of the sunspot cycle and i is an undetermined integer. Since the phase wandering of the sunspot cycle exceeds 2 years, i cannot be zero, so Dicke takes i = 1 to be the most plausible solution.

By contrast, a turbulent dynamo confined to the solar convection zone cannot be expected to maintain phase over long periods of time. Even though many theoretical idealizations of the dynamo, such as those based on mean-field electrodynamics, are described by equations that have periodic solutions, in reality one would expect turbulent fluctuations to destroy memory. Thus one might attempt to distinguish observationally between such dynamo hypotheses and the possibility of a regular oscillator by measuring the degree of phase maintenance of the sunspot cycle.

The first difficulty one encounters in such an endeavour is the problem of deciding how to define the phase of the cycle. Only the two most naive measures have been considered so far: the instant of field reversal which is estimated by the time of sunspot minimum, and the instant of greatest surface field which is estimated by the time of sunspot maximum.

Two independent analyses (5, 7) of the sunspot record have been carried out in an attempt to decide between the alternatives. Both used the same two statistical models to compare with the data, one with random fluctuations about a perfect clock and the other, which I shall misname the dynamo model, assumed random fluctuations in phase. The principle of both analyses was to choose a measure of the phase wandering of the cycle, and to compare the result with the expectations of the two models. The first discussion (7) was quite elementary, and used an obviously imperfect statistic that was chosen primarily for computational simplicity. The years of sunspot maxima and minima were used separately as tests, and it was found that the phase wandering of sunspot maxima lies closer to the expectation of the clock model, and that of sunspot minima is closer to the expectation of the turbulent dynamo model.* It was concluded, therefore, that the data is inadequate to support either model.

---

* Formulae (8.7) and (8.8) in ref. 7 were quoted incorrectly. The ratio of the expectation of the square of the phase deviations to that of the period fluctuations should be $N(N+1)/[6(N-1)]$ and $N(5N-1)/[6(N^2-1)]$ for Models A and B respectively. Correcting these results does not alter the conclusion.

The second analysis (5) was more sophisticated. It employed a statistic that would have been less biased than that in (7) had the raw sunspot data been used. However, the presumed correlation in Dicke's model between the rise time of the field from the core and the sunspot number at subsequent sunspot maximum was used to adjust the dates of sunspot maxima to move them as close as possible to the clock model. These dates were used also to test the dynamo model; and the dates of sunspot minima were not considered. It is perhaps not surprising, therefore, in view of the results of (7), that the sunspot data appeared to be in closer agreement with the expectation of the clock model. It was concluded that the data tends to support the clock model shows no statistical indication of random fluctuations in phase.

How confident can we be in this conclusion? I shall illustrate the use of the statistics in terms of the more elementary analysis of ref. 7. Consider a sequence of N successive sunspot cycles. For the clock model suppose that the time of occurrence of the nth maximum (or minimum) after the first (to which I assign n = 0) is $t_n = nT + \tau_n$, where T is constant and the τ_n are independent random variables with zero mean and standard deviation τ. The period of the n^{th} cycle is $P_n = T + \tau_n - \tau_{n-1}$, and the mean period of the N cycles is $\bar{P}_N = T + N^{-1}(\tau_N - \tau_0)$. For the dynamo model, assume that the interval between two successive cycles is $P_n = \Psi + \psi_n$, where Ψ is constant and the ψ_n are also random variables with zero mean, but this time with standard deviation ψ. This model is really an extreme representation of a dynamo because it assumes that the sun has no memory of previous cycles at all. In this case

$$t_n = n\Psi + \sum_{m=1}^{n} \psi_m, \qquad \bar{P}_N = \Psi + N^{-1} \sum_{m=1}^{N} \psi_m .$$

The object of the investigation is to compute a measure of the phase deviation from a perfect clock. In the context of the clock model, this is an obvious measure of τ. Of course we do not know which clock to choose, and for maximum simplicity I shall choose the clock that ticks at the average rate of the cycle, at times $T_n = n\bar{P}_N + \varepsilon$, where ε is a constant. A measure of the phase deviation is the variance σ_ϕ^2 of $\phi_n = t_n - T_n$, defined by

$$\sigma_\phi^2 = (N+1)^{-1} \sum_{n=0}^{N} \phi_n^2 - \left[(N+1)^{-1} \sum_{n=0}^{N} \phi_n\right]^2 , \tag{2.1}$$

which is independent of ε. This can be computed from the sunspot data, and its expectation can easily be evaluated for each model. The result is

$$E(\sigma_\phi^2) = \frac{(N-1)(5N-1)}{3N(N+1)}\,\tau^2, \quad \frac{N-1}{6}\,\psi^2 \tag{2.2}$$

for the clock and dynamo models respectively, irrespective of the forms of the probability distributions of the τ_n and ψ_n. As $N \to \infty$ the prediction of the clock model remains bounded, whereas the dynamo model predicts an increase without limit.

Of course we have only a limited amount of data at our disposal, but we could compare the predictions of the two models with the sunspot data by dividing the data into segments of N cycles and comparing the dependence of $\overline{\sigma_\phi^2}$ on N, where the overbar denotes an average over the segments. I shall not present the results of doing simply that, but instead I introduce a second statistic that relates more to the dynamo model. This is the variance of P_n, which is an obvious measure of ψ. It is given by

$$\sigma_p^2 = N^{-1} \sum_{n=1}^{N} (P_n - \bar{P}_N)^2, \tag{2.3}$$

and its expectations are

$$E(\sigma_p^2) = 2(1-N^{-2})\tau^2, \quad (1-N^{-1})\psi^2 \tag{2.4}$$

for the two models. One can now consider the ratio $R \equiv E(\sigma_\phi^2)/E(\sigma_p^2)$, which is independent of τ or ψ, and compare it with $S \equiv \overline{\sigma_\phi^2}/\overline{\sigma_p^2}$ computed from the sunspot data.

The analysis is confined to the interval from the sunspot maximum of 1705.5 until the last sunspot minimum. The values of t_n have been taken from Allen (8), except that the dates of the first maxima in the nineteenth and twentieth centuries were replaced by 1803.5 and 1906.0*, and the date of the last sunspot minimum was taken to be 1976.5. In Figure 1 is shown the result of dividing the 24 cycles into q contiguous groups of $N = 24\ q^{-1}$ cycles for $1 \leq q \leq 6$. The rhombuses represent the ratio of the mean variance $\overline{\sigma_\phi^2}$, (averaged over the q groups) to $\overline{\sigma_p^2}$ for sunspot maxima; and the squares are the ratios for sunspot minima. To give some idea of the scatter, the vertical lines show the standard deviation of σ_ϕ^2/σ_p^2. Shown also are the values of R for the two models.

In this analysis the raw dates of sunspot maxima and minima have been used. If, as is implied by Dicke's theory for example, there is a physical relation between the phase delay of sunspot maxima and sunspot number, this should be taken into account. Thus one can consider the modified time sequence:

$$\tilde{t}_n = t_n - r(R_n - \bar{R}_N) \tag{2.5}$$

derived from the times t_n of sunspot maxima, where R_n is the sunspot number at the n^{th} sunspot maximum, and $\bar{R}_N$ is the mean of R_n over the N+1 maxima. The

---

* The dates 1805.2 and 1907.0 are quoted by Allen (8). However, both these maxima are double (9, 10); the values for t_n used in the analysis here are better representations of the average dates, and are close to those used in ref. 7.

coefficient r was chosen by Dicke (5) such that σ_ϕ^2 was minimized for q = 1. Thus it is likely that this would reduce $\sigma_\phi^2/\sigma p^2$, at least when q = 1, and so move the data closer to the prediction of the clock model (lower curve). The results of analyzing the sequence $\hat{t}_n$ with Dicke's value of r is shown also in Figure 1. Even though the results are typically closer to the clock model, they can hardly be said to confirm it.

One is tempted to conjecture that the sun lies somewhere between the two models, having a memory of finite duration. If that is the case, how long is that memory? A step towards answering that question has been made by Barnes et al. (11, 12). They studied numerical simulations of a rectified oscillator that is randomly perturbed. They adjusted the bandwidth of the response to the perturbation such as to bring the variance of the fluctuations in period into agreement with the sunspot data, and found that with the same adjustment the variance in the simulated sunspot numbers at sunspot maximum also agreed with the real data. Moreover, the model produced intervals of about 50 years of continuous low sunspot activity, which occurred roughly once in 500 years. According to Barnes et al. the model has an inverse

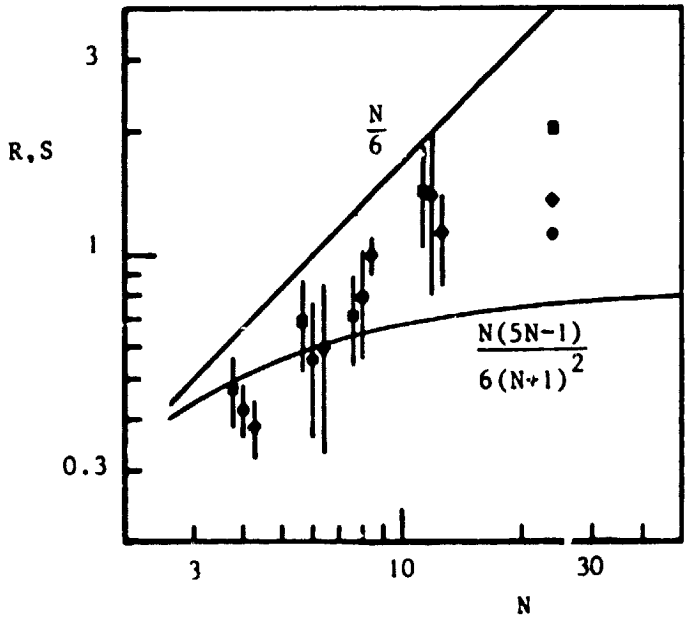


Figure 1. The ratios $S = \overline{\sigma_\phi^2}/\overline{\sigma p^2}$. The rhombuses represent sunspot maxima and the squares sunspot minima. Except when N = 24 they have been displaced horizontally to the right or left of the value of N to which they pertain, to prevent cluttering the diagram. The circles represent S for the time sequence $\hat{t}_n$ defined by equation (2.5) with r chosen to minimize σ_ϕ^2 for N = 24. The vertical lines extend to plus and minus one standard deviation of $\sigma_\phi^2/\sigma p^2$ from S. The continuous curves represent the ratios R of the expectations of σ_ϕ^2 and σp^2 derived from the two extreme statistical models.

bandwidth of 500 years, which is a measure of the memory, and an engineering rule-of-thumb is that many inverse bandwidths are required to establish whether or not phase is maintained. This suggests that with the direct observations that are available, it may not be possible to measure the sun's memory accurately.

POTENTIAL LUMINOSITY AND RADIUS VARIATIONS

It is quite reasonable to expect there to be small variations in the luminosity L and the radius R associated with the solar cycle. If the seat of the cycle is in the core, then any change in the size of the core would force the envelope to expand or contract, thereby modifying the hydrostatic stratification and hence R and L. The earliest photospheric response to any such change to the core occurs after a delay equal to the sound travel time from the core to the surface, which is about half an hour. Similarly, any change to the convection zone brought about by a turbulent dynamo would also produce modifications to the state of the photosphere. The question to which I now address myself is whether from the variations in L and R one might infer anything about the nature of the perturbation.

I shall first discuss in broad terms the sequence of events after an imaginary instantaneous perturbation to the solar structure, and then I shall discuss some specific examples in greater detail. I should point out straightaway that I do not have a definitive answer to the question, but the results of the discussion below are perhaps suggestive.

RESPONSE OF THE SUN TO AN INTERNAL DISTURBANCE

I have already pointed out that the fastest response to a perturbation is dynamical. The response to any large-scale perturbation that varies on a timescale of more than a few hours can therefore be regarded as being instantaneous and hydrostatic. I am not going to discuss dynamical oscillations here, and from now on I shall disregard the manner in which the relaxation to the new hydrostatic state takes place.

After hydrostatic adjustment follows thermal relaxation. There are three obvious thermal timescales outside the energy generating core that can be relevant to the evolution; these are quite disparate and therefore their manifestations can be discussed separately. The first is the time τ_a required for the convection itself to attain a balance with the mean stratification. This is of the order of the turnover time of the largest convective eddies. Taking the convection zone as a whole, this time is about a month, and measures the duration of the transient response to any deep-seated event. The equilibration time for the eddies near the surface, such as the granulation, is very much shorter. I shall be considering only changes that occur after times much greater than τ_a.

The second adjustment is the coming into balance of the radiation from the photosphere with the changed internal heat flux. This is what is

sometimes called the Kelvin-Helmholtz time for the convection zone. I shall call it τ_c. To within a dimensionless factor it is the ratio of the thermal energy in the convection zone to the solar luminosity. This ratio is about 10^5 years for so-called standard solar models. The dimensionless factor is probably of order unity, though recently it has been suggested that for the sun it is of order 10^{-4}. I shall return to this point later.

The third and longest time is the Kelvin-Helmholtz time for the entire sun. It is the ratio of the magnitude of the total energy of the sun, which by the virial theorem is approximately equal to the thermal energy, to the luminosity. It is also the thermal diffusion time τ_d characteristic of the entire sun, and is approximately 3×10^7 years.

Outside the core, the sequence of events following a perturbation is likely to be thus: after the convection zone has readjusted itself internally, on the timescale of a month, and the radiative interior has responded adiabatically, the entire convection zone either cools or heats up on a timescale of 10^5 years until a stratification is achieved with an essentially divergence-free heat flux. Finally, the radiative interior relaxes to its new state of thermal balance, on its thermal diffusion timescale τ_d. Notice that this is the sequence of events wherever in the sun the instigating perturbation may be located, though of course if that perturbation were confined to the superficial layers of the sun the magnitudes of the longer thermal responses may be imperceptibly small.

At this point I shall elaborate a little on what I have just said, in an attempt to dispel some common misconceptions about the meanings of these timescales. What I have to say is quite obvious to anyone who studies stellar evolution but does not appear to be common knowledge otherwise. The issue concerns whether the response of the photosphere to any deeply seated perturbation is evident in a timescale less than the Kelvin-Helmholtz time. I hope I have convinced you that in principle the answer must be yes, because the effect of any local change in the mass distribution of the sun will propagate with the sound speed. But suppose one considers a thermal perturbation somewhere in or at the base of the convection zone associated with which there is very little mass flow. In such a case there has been disagreement as to whether it is the thermal adjustment of the convection or the Kelvin-Helmholtz time for the zone that is important. This problem doesn't obviously arise when discussing the relaxation of the entire sun, because the analagous two times are the same. But surely the answer is this: both are important; the relaxation has two phases, and different processes control the evolution during the different phases.

In an efficient convection zone of any star, the convective adjustment time* τ_a is much less than the overall cooling time. Thus there is an initial internal redistribution of energy, on the timescale τ_a, followed by the slower evolution on the timescale τ_c which is controlled by the rate at which heat is radiated from the photosphere. This is the physics that describes evolution down the Hayashi track, for example, during which the structure of the entire fully-convective star is controlled by the radiation from the surface. It is analagous to the cooling of a hot block of copper: thermal conduction operates much faster than cooling from the surface, and the block is almost isothermal. It therefore cools at a rate that depends only on heat transfer processes at the surface and the thermal capacity of the block. The only essential differences in the case of a stellar convection zone are that the state of thermal balance is isentropic rather than isothermal and that the change in gravitational energy must be taken into account when assessing the thermal capacity of the convection zone.

In a radiative zone the evolution is quite different, for now it is the internal thermal readjustment that is the slowest. The analogy is now with the cooling of a block of wood. After an initial transient response during which the surface temperature adjusts to accommodate the heat flow from the interior, evolution proceeds on the thermal diffusion time, τ_d.

To summarize: a thermally relaxing convection zone adjusts its internal stratification in such a way as to supply the heat flow dictated by the surface conditions, whereas a radiative region adjusts its boundaries to transfer the heat that diffuses from within.

---

* In astronomy, this is often estimated as a thermal diffusion time obtained from a turbulent heat diffusivity computed from time-independent local mixing-length theory. If the action of the varying mean stratification on the dynamics of the turbulence is taken into account, still within the framework of local theory, the perturbation satisfies a wave equation instead, with a wave speed essentially equal to the rms convective velocity, w. The characteristic adjustment time of the entire convection zone is thus simply the advection timescale $\tau_w \equiv \int w^{-1} dr$, where r is a radial distance co-ordinate and the integral is over the vertical extent of the convection zone. It is likely that the dominant heat-carrying eddies in the main body of the convection zone actually extend from top to bottom. Thus aside from geometrical factors, τ_w is still a good estimate of τ_a, even though the local assumption is incorrect. Notice that the estimate of w via the relation $4\pi r^2 \rho w^3 \propto L$ (which follows from the usual considerations of the eddy dynamics that form part of mixing-length theory, but which does not depend on the detailed mixing-length assumptions about eddy breakup) is independent of the value of the assumed scale of the dominant eddies (or the mixing length) and is thus a fairly robust estimate. This estimate of w breaks down in the boundary layer at the top (and in the boundary layer at the bottom) of the convection zone, but that does not influence the value of the integral substantially.

COMPUTATION OF THE 'IMMEDIATE' RESPONSE TO A SPHERICALLY SYMMETRICAL PERTURBATION

Most of the numerical computations modelling the response of the sun on a timescale short compared with τ_c, but long compared with τ_a, have been integrations in time using a stellar evolution programme. The computations are quite expensive, and it is therefore worth contemplating other methods, even though they are more limited in scope.

If the disturbance is small everywhere, the obvious procedure is to perform a linearized perturbation analysis. I shall not discuss that here, simply because I do not yet have any results from this method. The reason is that my involvement in this problem stems from a search I once made for a non-linear relaxation oscillation involving the bottom boundary layer of the convection zone and its interaction with both the rest of the convection zone and the radiative region beneath. I was hoping to find a self-sustained oscillation in the solar luminosity with a period of about 10^5 years, with a view to explaining certain variations in the earth's climate. I therefore developed the following method, which treats the convection zone nonlinearly. My original investigation was never completed because I was unable to produce luminosity fluctuations with amplitudes greater than about 0.1 per cent, and at that time such changes were thought to be climatically insignificant.

The method involves only the computation of a few models of the convection zone, plus the linear adiabatic relaxation of the radiative interior. One is then able to estimate the change in the structure of the entire sun resulting from a given perturbation. Because one is seeking only the 'immediate' response, the total energy E of the star does not have time to change. Consequently, the idea is simply to compute the response of the model at constant E.

Let p_m and r_m be the pressure and radius of the model envelope at a fixed value of the mass co-ordinate m, whose value corresponds to the base of the convection zone in the unperturbed model. In addition consider a second envelope model with the same L and R into which a disturbance has been incorporated. Then if the pressure and radius at the same value of m in the second model are $p_m + \Delta p_m$, $r_m + \Delta r_m$, the actual change in p_m that would be produced by the same disturbance in a model of the entire sun is given by*

$$\delta \ln p_m \simeq \Delta \ln p_m + \left(\frac{\partial \ln p_m}{\partial \ln L} \right)_{R,A} \delta \ln L + \left(\frac{\partial \ln p_m}{\partial \ln R} \right)_{A,L} \delta \ln R, \qquad (3.1)$$

where $\delta \ln L$ and $\delta \ln R$ are the actual changes in $\ln L$ and $\ln R$. Similar relations give the variations in $\ln r_m$ and $\ln E$. Notice that for the linearized estimate (3.1) to be a good approximation it is no doubt necessary for the variations in E,L and R to be small, but it may be the case that in some limited regions there are properties of the envelope that vary by quite large amounts. Any localized nonlinearity that so arises is correctly taken into account.

The object now is simply to calculate $\delta \ln L$ and $\delta \ln R$ that result from the disturbance subject to the constraint $\delta \ln E = 0$. This can be done once the relation between $\delta \ln r_m$ and $\delta \ln p_m$ is known. To find that relation it is necessary to consider the response of the interior.

* The quantity A measures the amplitude of the disturbance. In practice the partial derivatives in equations (3.1) and (3.4) were evaluated at $A = 0$.

Because τ_d is large compared with the initial response time, the reaction of the radiative interior is essentially adiabatic. Moreover, since $\delta \ln p_m$ and $\delta \ln r_m$ are small, linearized theory may be used. The calculation is simply to solve the adiabatic radial pulsation problem at zero frequency in the radiative interior, without the mechanical boundary condition at the surface. This gives a relation between the pressure and radius perturbations on the radiative side of the base of the convection zone, which I represent by

$$\delta \ln r_m = \lambda \delta \ln p_m . \tag{3.2}$$

Since r_m and p_m must be continuous functions of m, condition (3.2) provides the information necessary for matching the perturbed convection zone onto its radiative interior, and together with $\delta \ln E = 0$ can be used to eliminate $\delta \ln r_m$ and $\delta \ln p_m$ from the relation (3.1) and its companions. The result is

$$\delta \ln L = \frac{D_R \, \Delta \ln E - (\Delta \ln r_m - \lambda \Delta \ln p_m) E_R}{D_L E_R - D_R E_L} , \tag{3.3}$$

with a similar equation for $\delta \ln R$, where

$$D_R \equiv \left(\frac{\partial \ln r_m}{\partial \ln R} \right)_{A,L} - \lambda \left(\frac{\partial \ln p_m}{\partial \ln R} \right)_{A,L}, \qquad E_R \equiv \left(\frac{\partial \ln E}{\partial \ln R} \right)_{A,L} , \tag{3.4}$$

and D_L and E_L are obtained by interchanging R and L. Notice that in general the procedure leads formally to a discontinuity in temperature at the base of the convection zone. In reality radiative diffusion and convective overshooting must smooth that out. If radiative diffusion alone were operating, after 10 years the jump in temperature would be spread over a layer only about 1000 km thick. Since this distance is only about one-third the mesh spacing near the base of the convection zone in my programme, the approximation is good. The artificial diffusion introduced by numerical differencing in a typical stellar evolution programme with a similar mesh spacing is likely to exceed greatly the radiative diffusion that is implied by the original differential equations.

It is also possible to estimate the relation between $\delta \ln L$ and $\delta \ln R$ that results from a disturbance that is confined to the solar core. Once again the expansion or contraction of the radiative envelope, outside the region of the disturbance, is adiabatic, and must match onto the perturbed envelope. If one were to assume that the perturbation could be linearized in the convection zone too, the problem would reduce to the nonadiabatic radial pulsation problem, with any period much greater than τ_a but much less than τ_c and without the boundary condition at $r = 0$. However, to get a rough idea of the result it is not necessary to solve that problem precisely. It is probably adequate simply to assume a homologous expansion or contraction of the convection zone, which yields equation (3.2) again, but with $\lambda = -0.25$. One can

then repeat the calculation, but with $\Delta \ln p_m = 0$, etc., to obtain

$$\frac{\delta \ln R}{\delta \ln L} = - \frac{D_L}{D_R}, \qquad (3.5)$$

where D_L and D_R are now computed with the new value of λ.

It is thus possible to estimate the response of the entire sun from numerical models of only the solar envelope and a knowledge of λ. The partial derivatives D_R, E_R etc. are computed by finite differences using undisturbed model envelopes with different values of L and R. Notice that although E, which is the energy of the entire star, can never be computed from envelope models alone, this does not matter because only differences of E appear in equation (3.3). These are simply the differences between the energies of the two appropriate envelopes above the matching point, plus the differences in the energies of the interiors. The latter can be computed as the work done by the envelope on the interior, which requires a knowledge of only p_m and r_m.

My computations reported below were performed with an early version of the computer programme used by Baker and myself (13) to model RR Lyrae pulsations. The programme had not been designed for this purpose, and several interpolations, which could have been avoided by rewriting the programme, were performed. Therefore I make no claims to high accuracy. The unperturbed model was chosen with abundances X = 0.745, Z = 0.02 of hydrogen and heavy elements, which are approximately the values that would have produced the correct luminosity in an evolved model of the entire sun. Cox-Stewart (14) opacities were used, and the equation of state was of the type discussed by Eggleton et al. (15). A mixing length of 2 pressure scale heights was chosen so as to yield a convection zone about 2×10^5 km deep, in accordance with the dictates of the high-degree five-minute oscillation data (16,17).

POTENTIAL INTERNAL CHANGES DURING THE SOLAR CYCLE

Convective inhibition by sunspots

The mechanism that has received most attention is the direct blocking of the heat flux by magnetic fields. This is particularly apparent in sunspots. Whether sunspots do actually reduce the energy flux has been questioned, the possibility being that the deficit in the radiative flux is made up by extra wave energy. The issue is probably not completely resolved, but I think the observational evidence is weighted towards a net reduction of energy flow within sunspots. I shall accept that here, and ignore wave transport entirely.

Notice that I have not yet said that the local reduction of the heat flow in sunspots necessarily implies a significant diminution in the solar luminosity. The local reduction is countered by a tendency for more heat to flow around the edges of the spot, producing a circumsistent bright ring.

However, the excess heat output by the identifiable part of the ring is quite inadequate to make up for the deficiency in the spot; inhibition of the flux in the spot extends deep into the convection zone,* so any excess flux can be distributed widely by the time it reaches the photosphere. Also any change in the mean (averaged over a spherical surface concentric with the photosphere) efficacy with which heat is transported in the convection zone leads to a change in the stratification of the convection zone that modifies the luminosity in such a sense as to oppose the original change. Nevertheless one expects that on average the opposing reaction is less than the perturbing influence, because the convection zone is apparently stable (though we are not absolutely sure of this). Thus it does appear that if the local inhibition of heat flow were the sole influence sunspots exerted, some reduction in luminosity would accompany an increase in sunspot numbers. But how substantial this reduction is cannot be judged without careful calculation.

It is very difficult to perform a realistic calculation to assess the effect of sunspot creation. What has been tried is to consider the effects of sunspots to be averaged over spherical surfaces r = constant, and to modify the standard techniques for studying spherically symmetrical stellar evolution to model the overall response of the sun. Thus, the magnetic inhibition of convective heat transport has been modelled by artificially reducing the mixing length ℓ in the usual time-independent heat flux formula. If such a procedure is a reasonable approximation to reality, it would be valid for studying variations on any timescale greater than τ_a, and would therefore be adequate for solar cycle variations.

The results of several independent computations have been published (18-22), and more are reported in this conference. The published results are summarized in Table 1, together with my own unpublished values. The diversity in the results arises partly from differences in the unperturbed models and partly from numerical error. For example, Dearborn and Newman (18) and Dearborn and Blake (21) used a small mixing length and consequently had a thin convection zone. In such a model the heat transport is more sensitive to the mixing length and it is therefore to be expected that $\delta \ln L/\delta \ln \alpha$ would be overestimated. To test my procedure I recently repeated the calculation with the ratio α of mixing length to pressure scale height equal to unity, yielding a convection zone with mass $9 \times 10^{-3} M_\odot$. Increased sensitivity was found, but my value of $\delta \ln L/\delta \ln \alpha$ was only 0.35, nearly a factor 2 less than that found by Dearborn and Newman (18) and Dearborn and Blake (21). I suspect that most of the scatter in the values of $\delta \ln R/\delta \ln \alpha$ and hence in

* By the argument in the antepenultimate footnote one expects the time taken for convection to adjust to the creation of a sunspot to be of order $\int w^{-1} ds$, where s is a distance co-ordinate and the integral is over a path that represents a typical heat-flow line that starts at the base of the spot and ends in the photosphere near the spot. The two independent observations reported at this conference that substantial transient reductions in the solar constant associated with large sunspot groups can last at least ten days therefore indicates that the spots responsible are not superficial phenomena.

TABLE 1 - Summary of published results from perturbing the mixing length

Authors	$\delta \ln L/\delta \ln \alpha$	$\delta \ln R/\delta \ln \alpha$	W	$M_c/M_\odot$
Dearborn and Newman (18)	0.46α †			0.008
Dearborn and Blake (21)	0.64	3×10^{-3}	5×10^{-3}	0.0065
Sofia et al. (19) Sofia and Endal (20)	1.5	0.11	7.5×10^{-2}	
Gilliland (22)			8.5×10^{-4}	
Gough	0.22	$\lesssim 1 \times 10^{-4}$	$\lesssim 5 \times 10^{-4}$	0.03

The last column is the mass of the convection zone in units of the solar mass.
† The formula for $\delta \ln L/\delta \ln \alpha$ quoted by Dearborn and Newman was found to hold for $1 \lesssim \alpha \lesssim 1.5$; the value of α corresponding to the quoted value of $M_c/M_\odot$ was not given. Dearborn and Blake used $\alpha = 1.4$.

$W \equiv \delta \ln R/\delta \ln L$ is a product of numerical error. In my calculations the numerator in the equation similar to (3.3) for $\delta \ln R$ is a small difference between two quantities of order unity (each of which is a numerical derivative computed from interpolated quantities). The result I obtain is sensitive to the interpolation formula I use, and I would therefore not be surprised if I have underestimated the degree of cancellation considerably. Thus at present I would summarize the results for a model with an adequately deep convection zone thus:

$$\frac{\delta \ln L}{\delta \ln \alpha} \simeq 0.3, \qquad W \equiv \frac{\delta \ln R}{\delta \ln L} \simeq 0. \tag{3.6}$$

In none of the publications is a relation between α and sunspot number derived. It is very difficult to do this theoretically, but one can be guided by observation. According to Allen (8) the mean intensity in a sunspot (umbra and penumbra combined) is about 70 per cent of that in the unperturbed photosphere. I shall take this as a measure of the inhibition of the convection at fixed entropy gradient. A medium-sized sunspot occupies about 6×10^{-6} of the surface of the sun, and at sunspot maximum there are present the equivalent of about 200 such sunspots.* Hence at sunspot maximum

* That is to say, about 100 on the side we can see and about 100 on the other.

the luminosity of the sun, ignoring faculae and the bright rings, is

$$L = (1 - \beta)\,\tilde{L}, \tag{3.7}$$

where $\tilde{L}$ is the luminosity the sun would have had had the sunspots been absent and the radiative flux been everywhere equal to the flux in the apparently undisturbed photosphere; β, which measures the mean reduction of the heat flux by the sunspots, is $200 \times (1 - 0.7) \times 6 \times 10^{-6} \simeq 4 \times 10^{-4}$. Throughout almost all the convection zone the convective heat flux is proportional to α^2, and consequently the observed mean flux inhibition is obtained by setting

$$2\,\frac{\delta\alpha}{\alpha} = -\,\beta. \tag{3.8}$$

We can now combine equation (3.8) with (3.6) to yield an estimate of the difference δL between the luminosities at sunspot maximum and sunspot minimum:

$$\frac{\delta L}{L} \simeq -\ 0.15\beta \simeq -6 \times 10^{-5}. \tag{3.9}$$

Notice that the actual decrease in the luminosity from sunspot minimum to sunspot maximum is only 15 percent of the apparent blocking of the luminosity $\beta\tilde{L}$ caused by the sunspots.

The exclusion of material from sunspots

It was stated by Jensen (23) that because the matter density in sunspots is lower than outside, the sun must have a larger volume at sunspot maximum. This idea has been elaborated on by Thomas (24) and Dearborn and Blake (25), who estimate the expansion to be not insubstantial. I do not understand their arguments, and I think it might be instructive if I say why.

Let us first consider the balance of energy in a star in hydrostatic equilibrium. In common with almost everyone else I shall ignore the turbulent stresses in the convection zone. I shall also assume that the pressure and the magnetic stress on the surface (i.e. the photosphere) can be ignored. Then the virial theorem takes the form

$$\tfrac{1}{2}\ddot{I} = 2T + \Omega + M, \tag{3.10}$$

where I is the so-called spherical moment of inertia, T, Ω and M are the kinetic, gravitational and magnetic energies, and the dots on I denote time derivatives. The kinetic energy T is comprised of the energy of thermal motion plus the energy of macroscopic motion, the latter residing mainly in

rotation. In equilibrium $\ddot{I} = 0$, whence

$$2T + \Omega + M = 0. \tag{3.11}$$

Suppose now that sunspots are created, and that at the same time the magnetic energy of the sun is increased by $\delta M > 0$. According to most theories of the sunspot cycle, the field is produced by the stretching that results from differential rotation. Thus the reaction of the field is to oppose gradients in angular velocity without changing the angular momentum, and hence to reduce the kinetic energy of the rotation. Just as the reduction of the luminosity by sunspots discussed above is not as great as the degree of direct inhibition of the heat flux, so the depletion of the rotational kinetic energy in this case is not as great as the energy imparted to the magnetic field. The continual driving of the large scale flow in the convection zone tends to restore the internal differential rotation to its original state, at the expense of other forms of energy. Let us assume, therefore, that only a fraction η (which I presume is positive) of the magnetic energy is extracted from T, and that the rest comes from Ω.* The virial balance is now

$$\begin{aligned} \tfrac{1}{2}\ddot{I} &= 2(T - \eta\delta M) + \left[\Omega - (1 - \eta)\delta M\right] + (M + \delta M) \\ &= -\eta\delta M < 0, \end{aligned} \tag{3.12}$$

so I must decrease. Hence, on average, the star shrinks, which might seem contrary to Jensen's assertion. It does not necessarily follow that there is a contradiction, however, for the adjustment might deviate substantially from being homologous. Nevertheless, it does illustrate a possible pitfall that might be encountered if one does not take into account that the magnetic energy in the sunspots must be provided from within the sun. Had I forgotten the changes in the other forms of energy in the star, as did Jensen and Thomas, I would have deduced that $\tfrac{1}{2}\ddot{I} = +\delta M > 0$, and then perhaps I might have been happy that this was apparently consistent with Jensen's claim.

Let us now look a little more carefully at the recent arguments. Thomas (24) evaluated the mass defect in toroidal flux ropes about 1000 km beneath the photosphere, and assumed that the total volume of the sun is simply increased by the volume that the missing mass would occupy at the ambient density. The estimated expansion was $\delta R/R \simeq 5 \times 10^{-4}$. Thomas's neglect of the change in the balance of forces in the interior consequent to the creation of flux ropes is tantamount to ignoring the fact that the region beneath the flux rope contracts as a reaction to the attempt to raise the height of the photosphere. Thus his estimates of the expansion must be exaggerated, and I shall now argue that the error may be quite large.

---

* Strictly speaking, the argument should be complicated further by considering also the interchange with the energies of ionization and the electrostatic interactions amongst electrons, ions and neutral atoms. These forms of energy do not appear in the virial theorem.

Suppose one simply implants a flux rope in pressure balance in the convection zone, and removes the mass defect ΔM from the star entirely. The perturbation to the gravitational field is negligible. Hence the stratification far from the flux tube is unaffected, and in particular the position of the photosphere is unchanged: one cannot detect a stationary submarine by observing the position of the surface of the ocean. Thus to compute the true effect of the flux tube on the photospheric radius one must merely replace the mass ΔM in the star. To within a factor of order unity, the resulting relative radius change will be $\Delta M/M_\odot$, and therefore the volume change is roughly equal to the volume occupied by ΔM at the mean solar density $\bar{\rho}$, rather than at the density ρ in the vicinity of the flux rope. If this argument is correct, Thomas has overestimated the expansion by a factor $\bar{\rho}/\rho$, which at 1000 km beneath the photosphere is about 5×10^5.

I am certainly not concluding from this exercise that the magnetic stresses of sunspots do actually have so miniscule an influence on the solar radius. All I am saying is that the exclusion mechanism discussed by Jensen and Thomas, taken in isolation, appears to be unimportant.

Dearborn and Blake (25) modelled the effect by including 'a global magnetic pressure term in a stellar structure code'. At first sight this appears to be the product of considering the sunspot magnetic field as providing an additional pressure contribution to influence directly the mean hydrostatic balance. But at this conference Dearborn has argued that his procedure can also be regarded as the excluded volume effect that Jensen and Thomas have discussed. He showed that if one integrates the stellar structure equations outside sunspots, where the magnetic field is negligible, then in order to relate the density ρ to dm/dr correctly one must add to ρ a term that takes account of the mass that has been pushed aside by the sunspots. This term is proportional to the magnetic pressure in the sunspots, and so appears as an additional contribution to the pressure in the equation of state. Precisely what this implies is difficult to judge, because we have not been told exactly how the additional term has been incorporated into the other equations. Dearborn and Blake find $\delta R/R \lesssim 10^{-4}$ associated with a 0.1 per cent reduction in the luminosity. Thus $|W| \lesssim 0.1$, and possibly $W \simeq 0$.

Other magnetic processes

Spiegel and Weiss (26) have considered recently the importance of the interaction between the convection zone and the radiative interior. They discussed the implications of the idea that magnetic field is compressed into a thin layer at the base of the convection zone by the combined action of topological pumping and field expulsion. After a sufficient amount of field has accumulated, hydromagnetic instabilities driven by magnetic buoyancy cause some of the field to rise to the surface to produce active regions and sunspots. Spiegel and Weiss suggest that the instability occurs when the magnetic layer is about a pressure scale height thick.

The cycle is complicated, and Spiegel and Weiss emphasize one aspect of it that may be important in causing luminosity and radius variations. It is that the magnetic layer will inhibit motion, in part by modifying the potential temperature gradient indirectly via the change in the hydrostatic

balance, and so cause the convection zone to recede. The flux expulsion process involves magnetic diffusion of stretched field that varies on a relatively short length scale. Since the magnetic Prandtl number is small, one might also expect thermal diffusion to be significant, so that the relatively quiescent magnetic layer would not be adiabatically stratified. But how it is stratified is hard to assess. A certain degree of mixing with the material in the radiative zone may well have taken place, so the resulting temperature gradient is presumably somewhere between the radiative and the adiabatic values.

An estimate of the manifestations of part of this process can be obtained from a somewhat different model I have constructed by suppressing the motion in a layer of thickness d at the bottom of the convection zone and assuming the reclaimed quiescent region to have achieved the radiative temperature stratification. Only one temperature discontinuity, at the base of the thermally mixed layer, was created. Magnetic stresses were not included in the hydrostatic balance in that region, so the calculation is not internally consistent. However, since the process being investigated is primarily thermal, it is not unreasonable to consider it in isolation from the balance of forces. Moreover, it is no less consistent than imagining α to have been changed without taking into account the stresses responsible. Indeed, the perturbation is equivalent to a drastic modification to α in a limited region of the convection zone. The result of a calculation with d = 7000 km is listed in Table 2. The increase in luminosity is about 5×10^{-4} $L_\odot$ somewhere near sunspot maximum, and is probably proportional to d^2.

TABLE 2 - Summary of the responses of the sun to various disturbances

Disturbance	$\delta \ln L$	$\delta \ln R$	W
Inhibition of convection by sunspots modelled by reducing α according to equation (3.8)	-6×10^{-5}	$\simeq 0$	$\simeq 0$
Introduction of equipartition tangled magnetic field into convection zone	2.6×10^{-2}	1.2×10^{-4}	4.4×10^{-3}
Replacement of temperature gradient in the bottom eighth of a pressure scale height of the convection zone by the radiative gradient	-4.5×10^{-4}	-9.1×10^{-5}	0.20
Any disturbance that is confined to the energy generating core			0.53

Although the magnitude of the luminosity perturbation is comparable with that deduced by Spiegel and Weiss, the mechanism by which it was obtained is rather different. Replacing what was essentially the adiabatic temperature gradient by the radiative gradient in the lower boundary layer of the convection zone results in a change in the stratification which is locally much greater than Spiegel and Weiss envisaged. Consequently, a greater redistribution of energy takes place. For example, subsequent to the magnetic instability, the temperature of the material near the base of the convection zone increases by about 10^4K and the internal energy in the convection zone is decreased by about 10^{42} erg. If this estimates the energy available to supply the increment δL in the luminosity, it would imply that $\delta L \approx 10^{-4} L_\odot$ if the increment were spread uniformly over 10^5 years. The interchange between the different forms of energy is brought about by a force provided from a comparatively small energy reservoir: to suppress the convection at the base of the zone and change the stratification to the radiative gradient requires only about 3×10^{33} erg of work. Thus magnetic energies as great as the total change in the energy radiated (if subsequently the star were to remain unperturbed until the relaxation time τ_c had elapsed) are not necessary to bring about that change. This result provides some a posteriori justification for ignoring the magnetic stresses in the hydrostatic equation.

In contrast, Spiegel and Weiss imagine the perturbation to cause a redistribution of the energy in the convection zone of just 10^{39} erg. This they assume is radiated in only about 10 years. The crux of the disagreement between their ideas and the assumptions of the calculation described above, therefore, is that Spiegel and Weiss assert that under these conditions real convection, unlike the predictions of the mixing-length formalism, reacts very sensitively to perturbations from beneath. This difference of opinion has little to do with the efficacy with which convection transports heat down a gradient of potential temperature. It concerns the degree to which convection modifies the photospheric temperature and so changes the rate at which heat is radiated from the star. The very high sensitivity of the state of the photosphere must be related to the small nonzero divergence of the heat flux in the convection zone, for if in the steady state the sensitivity of real convection were 10^4 times greater than the prediction of mixing-length theory, it would be unlikely that the latter could have been used to reproduce successfully the slope of the lower half of the main sequence in the Hertzsprung-Russell diagram.

Another phenomenon of interest is the influence of the small-scale tangled magnetic field in the convection zone. One effect is that the magnetic pressure modifies the hydrostatic balance, and another is that magnetic buoyancy enhances the driving force on the turbulent eddies and so increases the efficacy of the convection. I have modelled these processes by adding to the free energy of the fluid, from which all thermodynamical state variables are calculated, the energy of a tangled magnetic field in equipartition with the kinetic energy of convection. This increases the fluid pressure by an amount equal to the magnetic pressure. It also reduces the adiabatic temperature gradient and thus enhances the buoyancy forces acting on the convective eddies. As in the case of changing the mixing length to scale height ratio by a constant amount, this perturbation has a significant influence on the stratification only in the upper boundary layer of the convection zone. The idea is

that there will be more tangled field in the convection zone at sunspot maximum. Once again, to compare such a model with one having no magnetic field at all overestimates the difference between sunspot maximum and sunspot minimum. And indeed, the luminosity enhancement by the magnetic field at sunspot maximum of more than 2 per cent (see Table 2) is greater than the limits set by observation.

Perturbations to the core

Table 2 also contains an entry corresponding to the response to a perturbation that is confined to the core. The perturbation is presumed to provide only a mechanical disturbance to the base of the envelope. Thus no rising magnetic field of the kind envisaged by Dicke (1), for example, is accounted for. Without specifying the amplitude of the core perturbation one cannot set absolute values to the perturbations in L and R, but provided linear theory is valid, their ratio is independent of the nature of the perturbation.

CONCLUSION

We do not yet know whether the solar cycle is controlled in the convection zone or the radiative interior. It cannot be claimed that the sunspot statistics support either view convincingly, though they do hint that the sun does not keep perfect time. If that is indeed the case, one might regard it as evidence that a turbulent dynamo is operative, and that the wandering of the phase of the cycle is produced by the dynamical effect of the turbulent fluctuations on the oscillation. A convincing demonstration that the phase of the cycle is not maintained would not close the case, however, because it is quite common for nonlinear systems to oscillate almost but not exactly periodically without any stochastic interactions. The potential diagnostic power of the sunspot statistics lies mainly in the possibility of demonstrating phase maintenance, for in that case stochastic interactions must necessarily be unimportant.

Studies of the luminosity and radius variations associated with the cycle will probably be more fruitful. Some work has already been done, but mainly with only superficial perturbations meant to represent the magnetic inhibition of convection in the upper boundary layer of the convection zone. It may be that plausible variations in luminosity can be engineered, though the associated radius variations are very small: $W \approx 0$. The response of the sun to a few other types of disturbance have been discussed in this paper, but no systematic investigation has yet been undertaken. In all cases it is hard to estimate the absolute magnitudes of the resulting luminosity and radius perturbations, but their ratio W is more clearly determined. The examples suggest that W increases as the depth of the disturbance increases, and if that tendency is ever demonstrated to hold universally, it seems likely that imminent observations will enable us to decide at least whether part of the dynamo process operates deep in the sun.

Other diagnostics that might be of use in this respect come to mind. The low-degree five-minute oscillations provide integral measures of the solar

interior that must vary over the solar cycle. It has not yet been demonstrated, however, whethe they remain coherent for long enough to have sufficiently accurate frequencies to measure the solar change. Another indicator may be the apparent quasibiennial variation in the solar neutrino flux. Sakurai (27,28) has found a 26 month variation in the measurements of Davis and his colleagues which appears to be correlated with the residuals in the sunspot numbers that remain after subtraction of a 5-month running mean. If there is a causal connection between the variations of sunspots and the neutrino flux, its discovery would clearly be important. One conclusion we can draw straight away, however, is that none of the disturbances seated outside the core that have been considered here is of a magnitude anywhere near to being adequate to cause any perceptible variation in the neutrino flux.

I am very grateful to Dr N.O. Weiss for many interesting discussions.

REFERENCES

1. Dicke, R.H.: The clock inside the sun, New Scientist, vol.83, 1979, pp 12-14.

2. Lyttleton, R.A.: The Gold effect, Lying truths, ed. R. Duncan and M. Weston-Smith, 1979, Pergamon, Oxford, pp 181-198.

3. Cowling, T.G.: The present status of dynamo theory, Ann.Rev.Astron. Astrophys., vol. 19, 1981, in press.

4. Dicke, R.H.: The rotation of the sun, Stellar rotation, ed. A. Slettebak, 1970, Reidel, Dordrecht, pp 289-317.

5. Dicke, R.H.: Is there a chronometer hidden deep in the sun? Nature, vol. 276, 1978, pp 676-680.

6. Dicke, R.H.: Solar luminosity and sunspot cycle, Nature, vol. 280, 1979, pp 24-27.

7. Gough, D.O.: The significance of solar oscillations, Pleins feux sur la physique solaire, ed. J. Rösch, CNRS, Paris, 1978, pp 81-103.

8. Allen, C.W.: Astrophysical quantities, 3rd edition, 1973, Athlone Press, London.

9. Waldmeier, M: The sunspot-activity in the years 1610-1960, 1961, Schulthess, Zürich.

10. Eddy, J.A.: The Maunder minimum, Science, vol. 192, pp 1189-1202.

11. Barnes, J.A., Sargent III, H.H. and Tryon, P.V.: Sunspot cycle simulation using random noise, Geochim.Cosmochim.Acta, Suppl. 13, 1980, pp 159-163.

12. Barnes, H.A., Sargent III, H.H. and Tryon, P.V.: Sunspot cycle simulation using a narrowband Gaussian process, NBS Technical note 1022, 1980, US Gov. Printing Office, Washington.

13. Baker, N.H. and Gough, D.O.: Pulsations of model RR Lyrae stars, Astrophys.J., vol. 234, 1979, pp 232-244.

14. Cox, A.N. and Stewart, J.N.: Rosseland opacity tables for population I compositions, Astrophys.J., Suppl 19, 1970, pp 243-279.

15. Eggleton, P.P., Faulkner, J. and Flannery, B.P.: An approximate equation of state for stellar material, Astron.Astrophys., vol. 23, 1973, pp 325-330.

16. Lubow, S.H., Rhodes Jr, E.J. and Ulrich, R.K.: Five minute oscillations as a probe of the solar interior, Nonradial and nonlinear stellar pulsation, ed. H.A. Hill and W.A. Dziembowski, 1980, Springer, Heidelberg, pp 300-306.

17. Berthomieu, G., Cooper, A.J., Gough, D.O., Osaki, Y., Provost, J. and Rocca, A: Sensitivity of five minute eigenfrequencies to the structure of the sun, Nonradial and nonlinear stellar pulsation, ed. H.A. Hill and W.A. Dziembowski, 1980, Springer, Heidelberg, pp 307-312.

18. Dearborn, D.S.P. and Newman, M.J.: Efficiency of convection and time variation of the solar constant, Science, vol. 201, 1978, pp 150-151.

19. Sofia, S., O'Keefe, J., Lesh, J.R. and Endal, A.S.: Solar constant: constraints on possible variations derived from solar diameter measurements, Science, vol. 204, 1979, pp 1306-1308.

20. Sofia, S. and Endal, A.S.: Probable detection of climatically significant change of the solar constant, Geochim.Cosmochim.Acta, Suppl.13, 1980, pp 139-146.

21. Dearborn, D.S.P. and Blake, J.B.: Is the sun constant?, Astrophys.J., vol. 237, 1980, pp 616-619.

22. Gilliland, R.L.: Solar luminosity fluctuations, Nature, vol. 286, 1980, pp 838-839.

23. Jensen, E.: On tubes of magnetic force embedded in stellar material, Ann.d'Astrophys., vol. 18, 1955, pp 127-140.

24. Thomas, J.H.: Variation of the sun's radius and temperature due to magnetic buoyancy, Nature, vol. 280, 1979, pp 662-663.

25. Dearborn, D.S.P. and Blake, J.B.: Magnetic fields and the solar constant, Nature, vol. 287, 1980, pp 365-366.

26. Speigel, E.A. and Weiss, N.O.: Magnetic activity and variations in solar luminosity, Nature, vol. 287, 1980, pp 616-617.

27. Sakurai, K.: Quasi-biennial variation of the solar neutrino flux and solar activity, Nature, vol. 278, 1979, pp 146-148.

28. Sakurai, K.: Quasi-biennial periodicity in the solar neutrino flux: a further result, IPKU Research note 1001-80-1, 1980, Kanagawa University.

D20

SHORT AND LONG TERM VARIATIONS IN THE "SOLAR CONSTANT"

Kenneth H. Schatten
NASA/Goddard Space Flight Center

ABSTRACT

Short and long term variations in the solar constant are examined theoretically. The variations observed by the Solar Maximum Mission, lasting several days and associated with the passage of sunspot groups, strikingly demonstrates the well known lack of a "bright ring" effect around sunspots. This suggests that sunspot magnetic fields do not simply block the heat flowing upward into the photosphere. Rather, it is suggested that gravitational draining occurs; this cools sunspots and transports downward the heat that would otherwise flow into the photosphere. A model of sunspot temperature with depth shows modest support when compared with the empirical model of Van't Veer. Secular trends in the solar constant may occur and be associated with the influence of the convection zone magnetic field upon convective heat transport. As a start to understanding this problem, the Schwarzschild criterion has been modified to include the effects of magnetic field.

INTRODUCTION

Recently investigations by Livingston (ref. 1), Kosters and Murcray (ref. 2), Livingston et al. (ref. 3), Dicke (ref. 4), and Willson et al. (ref. 5) suggest that short term and secular changes in solar luminosity may be occurring. Theoretically, the influence of the solar cycle magnetic field on solar luminosity is a multifaceted question, since there are numerous ways in which magnetic field can affect the sun's luminosity.

Thomas (ref. 6) has theoretically investigated luminosity changes associated with effective solar radius changes produced by magnetic buoyancy in spots. Hoyt (ref. 7) looked at differences in luminosity associated with observable photospheric features-spots, faculae, etc. - to obtain a measure of possible global luminosity differences. He found a relationship between umbra-penumbra area ratios and terrestrial temperature variations. Willson et al. (ref. 5) have examined dips in the solar luminosity associated with the passage of spot groups past central meridian. These dips are a "short term" influence we shall discuss shortly. "Long term" influences may arise from changes in the convection zone. Schatten (ref. 8) suggested that magnetic buoyancy may influence convection; Spiegel and Weiss (ref. 9) argue that the heat transport by convection can be affected by the magnetic field, and that the influence is most significant at the base of the convection zone. We shall discuss, later, a mechanism that allows the long term influence of the magnetic field upon the convection zone to be calculated from stellar models. First, let us examine the short term variations associated with spot groups.

SHORT TERM VARIATIONS ASSOCIATED WITH SPOT GROUPS

As the photospheric material cannot serve as a reservoir for the vast flow of energy on time scales longer than about a second, a "bright ring" around sunspots might be expected if the Biermann (ref. 10) mechanism involving field inhibition of convective heat transport occurred.

A new school of thought developed, beginning with Danielson (ref. 11), in which the convective generation of Alfven waves below the sunspot consumed so much energy that cooling of the spot occurred with an attendant field concentration. It also eliminated the "bright ring" effect wherein the blocked heat would flow around the spot into the surrounding photosphere. The lack of a bright ring was strikingly demonstrated by the recent "solar constant" observations of Willson et al. where dips in the solar constant occurred when sunspots approached central meridian. If any bright ring were present, it would cancel the spot energy deficit and no dip should occur. It is these dips that are the subject of the present section. Other investigators have also found changes in the solar constant with sunspot visibility. Foukal and Vernazza (ref. 12) found a level of 3×10^{-4} change in the solar constant (similar to the SMM findings) and Chapman (ref. 13) found a dip in the sun's brightness equal to 62% of a sunspot's area.

Here we have followed much of the theoretical guidance suggested by Meyer et al. (ref. 14) and by Parker (ref. 15, 16). We further this third view that pores, knots and sunspots are a kind of dynamical solar sink in which material drains gravitationally downward. This utilizes a dynamical gravitational draining as the mechanism to explain 1) the cooling of features (ref. 15) 2) the elimination of the bright ring around sunspots, and 3) the field concentration mechanism (ref. 14). It enables a calculation of the temperature of these downdrafts to be roughly 1000°K cooler than the photosphere from basic principles and that the downflow in fluxtubes should be roughly several km/s.

Parker (ref. 16) has suggested that the "6-8 km s^{-1}... downdrafts in the fluxtubes (leading to) concentration to 1500 G may be a direct dynamical consequence". This downdraft inferred by Deubner (ref. 17) from observations may also be important in stabilizing fluxtubes insofar as the magnetic pressure of the fields, in a solar atmosphere with no inward and downward flow, would tend to disperse the fields in a short amount of time. Meyer et al. (ref. 14) discuss the stabilizing effect of this inflow of material as a hydrodynamic "collar" which the spot wears.

We discuss here, that this inward and downward flow as shown in Figure 1, may also play a role in the cooling mechanism for spots, as well as pores where Frazier (ref. 18) has observed downdrafts up to 3 km s^{-1}. For pores and knots, large velocities in the photosphere are found, or at least inferred however, for spots Beckers (ref. 19) notes that they "show no vertical motions ... exceeding 25m/sec." Some confusion has arisen "because of the limb effect of the surrounding photosphere, sunspots appear to have a downflow of ~400 m/sec ...." Thus downflows are seen or inferred for small flux tubes, and not for spots where we hypothesize a deeper gravitational draining occurs owing to the larger size of the magnetic object. In this view, the field

originates from dynamo processes, the final stage of which is magnetic buoyancy, whereby a large concentrated field erupts at the sun's surface. In the absence of a suitable cooling and field stabilization mechanism, the strong field would quickly dissipate. However, as the material near the sun's surface is continually radiating luminous energy into space, the gases cool and with their increased density return in conduits to the heat source from which they came, thereby completing the convection cycle. Thus by analogy with any thermal cycle, the material moving towards the heat source would be the coolest. In the absence of sunspot fields, the cooled photospheric gases return at the boundaries of the supergranulation, and so form pores. In the presence of sunspot fields, they are aided in their return by this field conduit, which tends to reduce the turbulent viscosity and so provides an easier pathway. Owing to the larger size of a spot, not enough material appears capable of congregating into an observable downflow in the photosphere and we hypothesize it does so at depth.

To calculate the size of the effect, we take a volume element within the magnetic region, as shown in Figure 1. It is bounded by 4 sides (1-4) with no interaction on the 5^{th} and 6^{th} sides due to azimuthal symmetry. We shall simplify the picture by assuming the following. The flow is roughly inward to the box at 1, and outward at 4. If the photosphere were not yet cool, region 2 would radiate the same energy rate, F, as the remainder of the photosphere. Due to the inward and downward flow, little or no convective energy is being transported into the region through 3 (where in the normal photosphere, a rate, F, balances the outflow through the top). For the purpose of the calculations, the volume extends from the photosphere $\tau = 1$ to roughly $\tau = 3$ so that most of the radiant flux, $F = 6.4 \times 10^{10}$ erg $cm^{-2}s^{-1}$ (ref. 20), is indeed emitted from the volume through 2. Further, the material flows inward through 1 at several km sec^{-1} and out through 2 at the same rate and pressure, so that we can ignore compressive heating or expansive cooling as a consideration. Then the temperature difference can be shown to be:

$$\Delta T = \frac{2}{5} \frac{FL}{\bar{\rho}\, hvR} \approx 1{,}100^{\circ}K \qquad (1)$$

where R is the gas constant, L and h are the length and height of the volume, taking $L/h \approx 1$ and $v = 7$ km s^{-1}. Thus for a spot, an umbral temperature of $4{,}600^{\circ}K$ is calculated showing good agreement with observed umbral temperatures. Put simply, the mechanism can provide adequate cooling for pores, magnetic knots and sunspot umbra. We now examine the question of the downflow. The cooled photospheric gases, being denser than the hotter photospheric gases, would tend to descend. Will this gas reach a terminal velocity of several km sec^{-1} as it descends? We consider a calculation for the terminal velocity of a blob of gas in a viscous stratified atmosphere.

As the terminal velocity is governed by the interaction with the stratified medium, we choose the scale height, H, to be a characteristic length dimension for consideration. The downward force upon the gas is:

$$F = g\, \Delta\rho\, H^3 \qquad (2)$$

where g is the acceleration of gravity $= 2.7 \times 10^4$ cm s^{-2}, $\Delta\rho$ is the

additional density of the blob of gas $\approx \frac{\Delta T}{T} \rho = \frac{1100^{\circ}K}{5700^{\circ}K} 4 \times 10^{-7}$ gm $cm^{-3} = 0.7 \times 10^{-7}$ gm cm^{-3}, and H is the scale height ≈ 100 km. The viscous balancing upward force is:

$$F = \frac{\eta v A}{d} \approx \eta v H \qquad (3)$$

where η is dynamic viscosity, $\rho\nu$, and ν is the kinematic viscosity $1/10$ vl. Busse (ref. 21) obtained a value of 10^{12} cm^2 s^{-1} for ν, yielding 3×10^5 gm $cm^{-1}s^{-1}$ for η. The quantity v is the velocity difference between the blob and the surrounding material, and H is the scale height of the photosphere again. Equations 2 and 3 can be solved for v, yielding:

$$v = g\Delta\rho H^2/\eta. \qquad (4)$$

This gives 7 km s^{-1} for the terminal downward velocity of gases, a sufficient downflow to provide cooling. It is comparable with the 2 km s^{-1} velocity that Parker (ref. 15) required for a downflow within spots to account for their cooling.

The lack of a bright ring around the gas central to the problem of solar constant dips is understood in this model by examining Figure 1. As the upward heat flux, F, transported principally by convection, encounters the sunspot magnetic field, the energy and gas are entrained amongst the descending gases where the energy is carried down toward the base of the convection zone and so is lost to the sunspot umbra.

Table 1 shows the temperature and density of the umbra and surrounding photospheric material. The photospheric parameters are shown from the solar model of Endal and Sofia (ref. 22). The umbral parameters were obtained from a computer model having been calculated utilizing the simple theoretical model outlined. That is, there is no convective heat transport into the spot from below and the photospheric material is cooled by radiation into space.

For comparison, the semi-empirical model of Van't Veer as outlined in Tandberg-Hanssen (ref. 23) is shown in table 2. It is based on a method developed by Van't Veer that determines the umbral parameters using the measured intensity in the wings of certain Fraunhofer lines. Near $\tau = 4$, the values agree remarkably well; however, near $\tau = 1$ the agreement is less good.

LONG TERM VARIATIONS IN THE SOLAR CONSTANT

We consider, in this section, the effects of the solar cycle magnetic field embedded within the convection zone upon the solar luminosity. We develop a version of the Schwarzschild criterion which includes the effects of magnetic field. We consider the approximation that the magnetic field provides a net isotropic pressure.

The Schwarzschild criterion (ref. 24) is obtained by considering density changes associated with a rising convective element of the star, leading to instability if the radiative temperature gradient exceeds the adiabatic gradient:

$$\left(\frac{d \ln T}{d \ln P}\right)_r > \left(\frac{d \ln T}{d \ln P}\right)_a = \frac{\gamma - 1}{\gamma} \qquad (5)$$

If we now include a magnetic field, there will be a magnetic pressure P_B, associated with a rising bubble. In any discussion of the magnetic stresses, we may neglect the tension term in the stress tensor because the field is continuous and divergence B equals zero. Thus across any small element, the field tension balances on either side and imparts no force to the gas. The total pressure P_T, equals the gas pressure plus magnetic pressure:

$$P_T = P_G + P_M = P_G\,(1 + \beta^{-1}) \tag{6}$$

where β is the gas to magnetic pressure ratio. The adiabatic density gradient now becomes:

$$\frac{d\ \ln \rho}{dr}\Big)_a = \frac{d\ \ln P}{\gamma\ dr}\Big)_a - \frac{d\ \ln(1+\beta^{-1})}{\gamma\ dr}\Big)_a \tag{7}$$

where we have used the adiabatic relation, $P_G V^{\gamma}$ = constant. The radiative gradient is similarly modified:

$$\frac{d\ \ln \rho}{dr}\Big)_r = \frac{d\ \ln P_T}{dr}\Big)_r - \frac{d\ \ln T}{dr}\Big)_r - \frac{d\ \ln(1+\beta^{-1})}{dr}\Big)_r \tag{8}$$

Instability occurs if the adiabatic density gradient has a larger magnitude than the radiative gradient. From equations 7 and 8 this implies:

$$\frac{d\ \ln T}{d\ \ln P_T}\Big)_r > \frac{\gamma-1}{\gamma} + \frac{d\ \ln(1+\beta_a^{-1})}{\gamma\ d\ \ln P_T} - \frac{d\ln(1+\beta_r^{-1})}{d\ \ln P_T} \tag{9}$$

where adiabatic and radiative subscripts on β allow for differing values between convective bubbles and the surrounding environment. Equation 9, our modified Schwarzschild criterion, can be rewritten as:

$$\frac{d\ \ln T}{d\ \ln P_T}\Big)_r > \frac{\gamma-1}{\gamma}\left[1 - \frac{d\ \ln(1+\beta_a^{-1})}{\delta\ \ d\ \ln P_T}\right] \tag{10}$$

where we have incorporated the two expressions on the right of 9 into one through the utilization of a parameter, δ. Here δ equals one under the assumed approximation, for the onset of instability. For the more general case δ depends on the ratio of field inside the rising gas bubble to field outside, and geometry. Other formulae can be developed to include the effects of other geometries (see Thomas and Nye (ref. 25)). Further, if one excludes the interaction of the bubble with the material, an adiabatic equation of state can be written, but we have allowed this to be incorporated into δ. As we have developed our criterion for the onset of instability when the field is not perturbed, we have ignored the possible restoring force (or tension) between the base field and a rising bubble which may ensue with turbulence. The above criterion assumes that once the magnetic field and temperature gradient are suitable for instability, the Rayleigh-Taylor instability or another plasma interchange instability will develop to allow the bubble to rise.

The term on the right of equation 10 modifies the adiabatic condition in the following fashion. Let us consider β^{-1} to approach 1 in a region of the sun and examine how this affects the onset of instability. Figure 2 shows the magnetic field, the magnetic to gas pressure ratio (β^{-1}), and the term in equation 10 modifying the Schwarzschild criterion. As the gas pressure decreases radially outward, the whole term in the square bracket will first exceed one, and then fall below one, to possibly a negative value. Thus a single layer or region of magnetic field will form a relatively stable layer in its lower side and a relatively unstable layer in its upper side. The word

relative is used to indicate that these layers may be stable or unstable, depending upon the remainder of the Schwarzschild condition, and thus, they are merely modifying the condition. However, in case the modifying term reaches a negative value, as the left hand side of equation 10 is positive definite, instability is required. This is conventionally referred to as magnetic buoyancy, and in the usual treatment only the field considerations are provided. We see here, however, that ordinary convection may also be affected by the presence of a subsurface field.

If we consider the fields in the solar interior suggested by dynamo theory (ref. 26), the term on the right of equation 10 appears to be significant only at the base of the convection zone (except in centers of activity). If the gas pressure variation has a scale height H_G, and the magnetic pressure a scale height H_B, the term on the right can reach a magnitude of order H_B/H_G. One of the problems associated with solar dynamo theory is the formation of a region where magnetic fields may regenerate without being lost to magnetic buoyancy (ref. 26, 27). In this result the underside of a magnetic field region would be a stable location where fields could regenerate. Above the region, if the field approached a high enough value, instabilities could form with a balance developing between regeneration and deterioration. Unno and Ribes (ref. 28) have examined magnetic buoyancy and have found that 200-300 gauss dynamo field can be retained in the convection zone for time scales up to about 80 years due to turbulent viscosity. The large scale field, although retained, could affect the onset of convective instability through the above criterion, thus the process of magnetic stabilization and destabilization may influence heat transfer at the base of the convection zone with solar cycle phase much as Spiegel and Weiss (ref. 9) suggest.

CONCLUSIONS

Downflow of material within sunspots is seen as a key to the understanding of their stability, temperature and heat flow. The inward velocity prevents the magnetic field from expanding into the surrounding surface by balancing the spot's magnetic pressure against the velocity change from an inflow into a downflow. The cool, dark characteristics of sunspots are to be explained by their being the cool end of a thermal cycle. In this cycle, the gases have radiated their energy into space, are allowed to descend along field conduits into a cooler, denser state and cut off the heat flow. The heat flux that would normally be convected into and radiated by the sunspot is carried downward by the flow to deep layers in the convection zone. Thus, the minima in the solar output associated with sunspot groups observed by Willson et al. (ref. 5) are explicable.

Further, for secular changes, we consider the magnetic field at the base of the convection zone to affect the Schwarzschild criterion and form stabilizing and destabilizing layers that may cause a significant variation in the solar luminosity similar to that in the model of Spiegel and Weiss.

REFERENCES

1. Livingston, W. C.: 1978, Nature, 272, 340.
2. Kosters, J. J., and Murcray, D. G. 1979: Geophy. Res. Lett., 6, 382.
3. Livingston, W., Milkey, R. and Slaughter, C.: 1977, Astrophys. J., 211, 281.
4. Dicke, R. H.: 1978, Nature, 276, 676.
5. Willson, R. C., Gullkis, S., Janssen, M., Hudson, H. S., and Chapman, G. A.: Science, 211, 700, 1981.
6. Thomas, J. H.: 1979, Nature, 280, 662.
7. Hoyt, D. V.: 1979, Climate Change, 2, 79.
8. Schatten, K. H.: 1973, Solar Physics, 33, 305.
9. Spiegel, E. A. and Weiss, N. O.: 1980, Nature, 287, 616.
10. Biermann, L.: 1941, Vierteljahrsschr. Astron. Ges., 76, 194.
11. Danielson, R. E.: 1962, Astron. J., 67, 574.
12. Foukal, P. V., and Vernazza, J. E.: 1979, Astrophys. J., 234, 707.
13. Chapman, G. A.: 1980, Astrophys. J. Lett., 242, 245.
14. Meyer, F., Schmidt, H. U., Weiss, N. O., and Wilson, R. P.: 1974, Mon. Not. R. Astr. Soc., 169, 35.
15. Parker, E. N.: 1979, Astrophys. J., 204, 259.
16. Parker, E. N.: 1979, Cosmical Magnetic Fields, Clarendon Press, Oxford, p. 213, p. 255.
17. Deubner, F.: 1976, Astron. and Astrophys., 47, 475.
18. Frazier, E. N.: 1970, Solar Physics, 14, 89.
19. Beckers, J. M.: 1981, The Sun as a Star, Ed. S. Jordan, NASA/CNRS, GPO, CH. 2.
20. Allen, C. W.: 1963, Astrophysical Quantities, Athlone Press, University of London, Ch. 9.
21. Busse, F. H.: 1969, Astrophys. J., 159, 629.

22. Endal, A. S. and Sofia, S.: 1981, Astrophys. J., in press.

23. Tandberg-Hanssen, E.: 1966, Solar Activity, Blaisdell Pub. Co., Waltham, p. 199.

24. Schwarzschild, K.: 1906, Gottinger Nachr., 41.

25. Thomas, J. H. and Nye, A. H.: 1975, Physics of Fluids, 18, 490.

26. Parker, E. N.: 1976, in Basic Mechanisms of Solar Activity, Ed. Bumba and Kleczek, 3.

27. Parker, E. N.: 1975, Astrophys. J., 198, 205.

28. Unno, W. and Ribes, E.: 1976, Astrophys. J., 208, 222.

Table 1

Depth km	Density 10^{-7}gm cm^{-3}	T_p 10^{3o}K	T_u 10^{3o}K
0	1.69	5.75	4.60
42	1.94	6.46	5.04
98	2.08	8.17	5.28
238	2.98	10.13	6.94
448	5.18	11.57	9.73

Table 2

τ(5000)	Depth, km from τ=.5	T_p 10^{3o}K	T_u 10^{3o}K
.01	-220	4.70	3.57
.05	-135	4.93	3.70
.1	-100	5.07	3.78
.3	-50	5.51	4.02
.5	0	5.83	4.18
1.0	40	6.41	4.48
2.0	60	7.18	4.84
4.0	100	8.10	5.24

CROSS SECTION OF A SUNSPOT

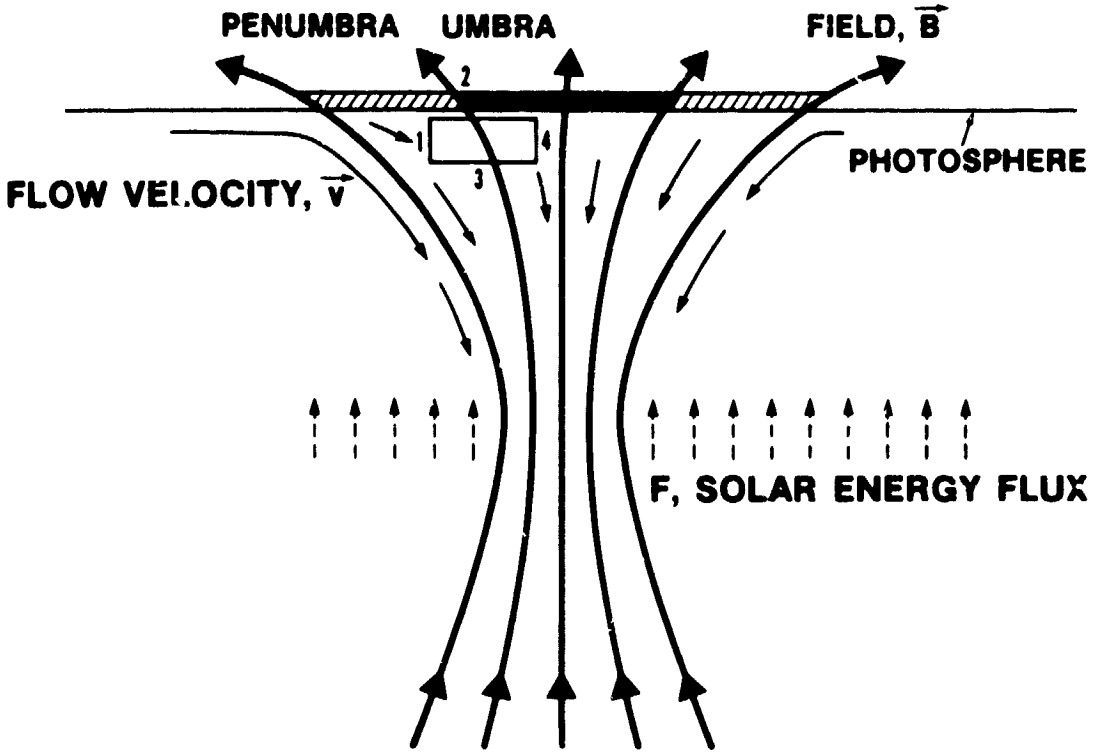


Figure 1. Shown is the field geometry and downflow velocities in the vicinity of a sunspot or pore. The inward and downward flow beneath the photosphere and below acts as a "collar," stabilizing the sunspot. This flow also returns the cooled gases to the base of the convection zone. It is by the liberation of radiant heat that the gases are cooled to umbral temperatures. The energy flux that would flow upward into the sunspot is entrained within the cooler gases and descends to the convection zone base to be reheated. The box at the top labelled with sides, 1-4, exhibits a simple heat calculation for th sunspot.

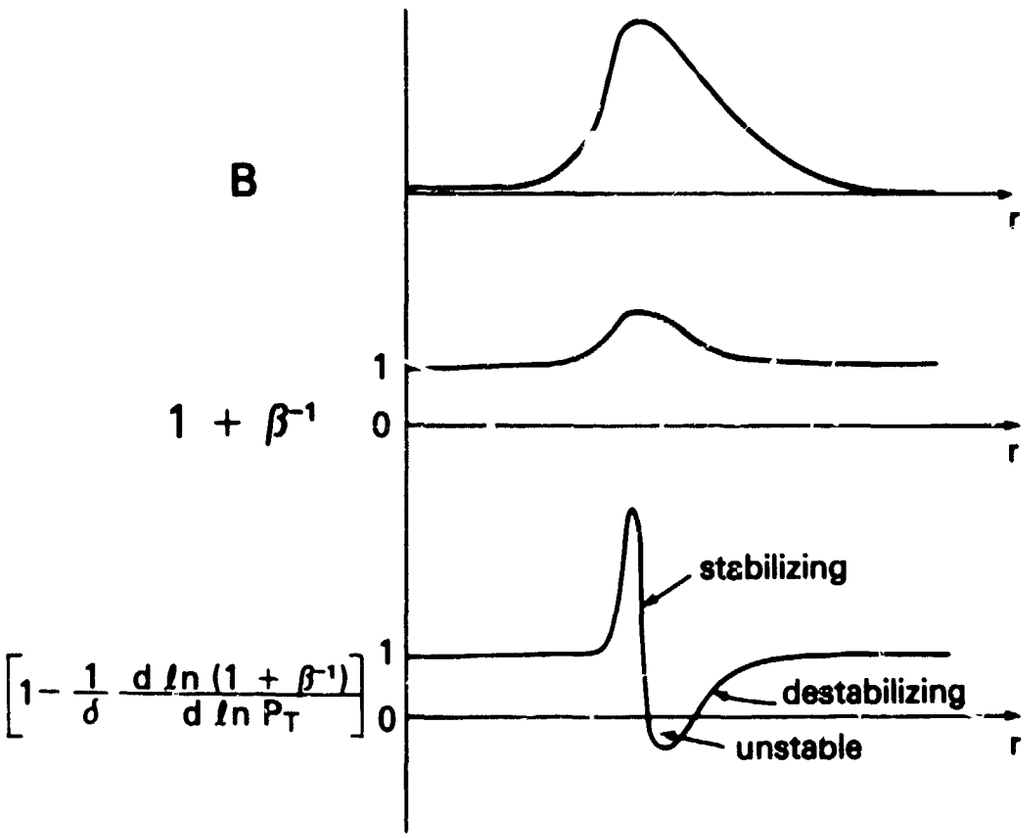


Figure 2. Schematic showing the radial variation of magnetic field B, magnetic to gas pressure ratio β^{-1}, and the term modifying the Schwarzschild criterion. As can be seen when the modifying term exceeds 1, the instability is more difficult to attain as convective elements have difficulty rising through a more tenuous field saturated layer. On the upper side of a field layer, however, destabilization occurs, and convective elements may more easily rise. When the modifying term becomes negative, then absolute buoyancy occurs as the field strength is sufficient to cause instability even against an unfavorable temperature gradient.

D21

MONITORING SOLAR-TYPE STARS

G. W. Lockwood
Lowell Observatory

ABSTRACT

Old UBV and recent uvby photometry of solar-type dwarfs and other standard stars yield an upper limit of variability (determined by observational errors) of about 0.004 mag rms. A factor two improvement in this upper limit is achievable.

INTRODUCTION

Current measurements of the total solar flux have a short-term precision of ∿0.1 percent and promise to yield a long-term accuracy better than 1 percent. Small variations of the total solar flux have been detected on a time scale of days by the Solar Maximum Mission (see elsewhere in these proceedings), and a possible change of the solar spectral irradiance in the visible region on a longer time scale may have occurred (ref. 1). In both cases, the change amounted to less than 0.5 percent, so in addressing the question of luminosity variation of solar-type stars, we must seek out photometric data with extremely high long-term accuracy. Absolute photometry of starlight in the visible and near-infrared approaches an accuracy of 1 percent only with great difficulty (ref. 2), but relative photometry, in which one star is compared with other(s) nearby in the sky is routinely done with a precision of ∿0.3 percent over time scales of months. Maintaining this precision over a decade is considerably more problematical but is feasible.

It is probably safe to state that the variations of solar-type main-sequence dwarf stars do not exceed ∿1%, because if larger variations were common, they would undoubtedly have been detected during the extensive photometric surveys of the past 30 years. A number of these stars have, in fact, been used as photometric standards.

PHOTOMETRIC STUDIES

BV PHOTOMETRY OF SOLAR-TYPE STARS

Jerzykiewicz and Serkowski published in 1966 (ref. 3) the results of a twelve-year photometric study in which 16 F- and G-type dwarf stars were

observed about ten times per year in the blue (B) and yellow (V) passbands of the UBV system. *To my knowledge, this program was the only long-term photometric study specifically addressed to the question of variability of solar-type stars.* These stars served as their own standards, thus eliminating errors caused by making transformations to an independent network of UBV standards. However, the stars were spread over seven hours of right ascension and were often observed at high airmasses; consequently, various systematic errors are likely to be present at some significant level.

A substantial effort was made to maintain the long-term stability of the measurements, but "improvements" in instrumentation were inevitable. In particular, a variety of photomultipliers was used, sometimes cooled by dry ice and sometimes operated at ambient temperature. The photometry of standard stars, Uranus, and Neptune has subsequently been analyzed for systematic effects related to changes in the instrumentation or the brightness and colors of the various objects (ref. 4). None could be found, although there was a systematic monotonic drift in one of the instrumental transformation coefficients during the course of the program.

Jerzykiewicz and Serkowski concluded that ". . . for none of these stars does the standard deviation of the yearly mean magnitude exceed 0.008 and for the stars 40 Leo, β CVn, and η Boo, this deviation is less than 0.004 mag." As shown in figure 1, taken from ref. 3, the peak-to-peak range of these stars is ∿0.01-0.02 mag.

What is the likelihood that the observed variations are intrinsic? Simple statistical tests provide little guidance: For example, a closer look at the published nightly magnitudes of individual stars reveals that there were very few statistical outliers. Hence, the fluctuations of the annual mean magnitudes cannot be substantially reduced by rejecting individual bad data points here and there. Furthermore, inspection of figure 1 reveals little correlation between one star and another, as would be expected in the case of systematic instrumental difficulties.

The internal night-to-night consistency of the magnitudes of individual stars improved dramatically after 1961, presumably owing to improvements in equipment or technique. This is illustrated in figure 2, where we have replotted the annual mean magnitudes (shown in figure 1) for two typical stars, ρ Gem (F0V) and 61 Vir (G6V), in order to include error bars indicating the standard errors of the annual mean magnitudes. The average number of observations per year was 8 from 1955 to 1961 and 11 from 1962 to 1966; hence, the statistical weights of the annual data points are comparable. The standard error of a single observation was 0.015 mag prior to 1961 and 0.007 mag afterwards.

Figure 3a is a histogram of the standard deviations of the annual mean V magnitudes of these 16 stars for the years 1962-1966. The average standard deviation is 0.004 mag, and the largest standard deviation is 0.006 mag. The average peak-to-peak variation (from figure 1) is 0.009 mag.

These values certainly provide a reliable upper limit for the variability of this sample of stars over this time interval. The weight of largely circumstantial evidence (viz., the changes in instrumentation, the frequent observations at high airmass) and the absence of a set of "control" stars showing smaller annual fluctuations than solar-type dwarfs argue strongly against a positive detection of variability. Indeed, the authors themselves concluded that ". . . no evidence of variability . . . has been detected . . ."

RECENT RESULTS

Beginning in the late 1960s, two developments promised an improvement in the ultimate accuracy of ground-based photoelectric photometry. Direct-current recording techniques were largely replaced by photon-counting methods, thus largely eliminating amplifier gain calibration and linearity problems. Intermediate-band uvby photometry utilizing interference filters offered relative freedom from the substantial photomultiplier-dependent color effects present in the broadband UBV photometric system. (These were especially troublesome in the V passband, whose shape on the long-wavelength side of the maximum was defined solely by the sensitivity of the photomultiplier.)

Unfortunately, no study comparable in scope and longevity to the one described above has been conducted using the uvby system. To proceed further, we consider some measurements obtained for other purposes using, coincidentally, the same 0.4-meter telescope used for the previous program. The incidental nature of these observations seriously compromises their usefulness for the purpose we now, belatedly, assign to them. Nonetheless, they are worth examining, and we shall demonstrate that modern equipment has allowed us largely to duplicate the result of Jerzykiewicz and Serkowski under otherwise unfavorable circumstances.

Since 1972, a set of ∿60 stars, including by chance 14 F-, G-, and early K-type dwarfs, has been used as standard stars for b (472 nm) and y (551 nm) photometry of planets and satellites at the Lowell Observatory (ref. 1, 5). Because these stars were used as standards, they were observed at a variety of airmasses and at all seasons of the year. But, most importantly, from early 1974 on, there were no changes in the photometer, the electronics, or the telescope, all of which were reserved solely for this particular program.

To achieve reasonable statistics, we have divided the data into three-year bins (1972-1974, 1975-1977, 1978-1980) and compared the range of variability of the 14 solar-type dwarfs to 14 other standard stars (giants and early-type dwarfs) for which comparable data were available. Figures 3b and 3c show, respectively, the distribution of peak-to-peak fluctuations seen in the three-year averages observed in these two sets of stars. For both groups, the average range of the three-year mean magnitudes was 0.004 mag. Figure 4 shows the amplitudes as a function of spectral type.

The two distributions do not differ significantly in shape, leading us to conclude that in both cases the observed "variations" are simply errors of observations.

Observers of variable stars routinely obtain differential magnitudes of pairs of stars with a short-term (days) precision of 0.002 mag or better. The following example demonstrates the accuracy attainable under tightly controlled conditions for a small group of stars near one another in the sky.

As part of the Lowell program described above, several groups of planetary comparison stars have been observed since 1972 (ref. 5). These are field F-, G-, and early K-type stars of unknown luminosity class selected for similarity of color and magnitude to the corresponding planet. As an example, we consider the Neptune comparison stars, a group of 16 stars in the magnitude range $6.7 < \underline{V} < 9.2$ located on a $\sim 15°$ arc of the ecliptic at declination $\sim -22°$. The typical number of observations per year was ~ 6. All 16 stars were observed near transit each night during the course of about an hour at an average airmass ~ 1.8. Extinction and time-dependent instrumental effects are thus minimized; and, although the average airmass was high, the differential airmass between one star and the others was usually quite small. After reduction to an ensemble mean, the average standard deviation of the annual mean magnitudes was 0.003, and the largest was 0.005 (figure 3d).

On the basis of this experience, it is reasonable to expect that a set of solar-type stars passing near the zenith could be monitored with an accuracy approaching 0.002 mag over a similar length of time. We estimate that a ten-year study involving ~ 25 stars could be carried out at a cost of \$500K (1980 dollars) using conventional ground-based, non-automated facilities. With an automated telescope, the data rate could be increased by about a factor of two at a dollar cost of an order of magnitude higher.

REFERENCES

1. Lockwood, G. W., Thompson, D. T., and Lumme, K.: A Possible Detection of Solar Variability from Photometry of Io, Europa, Callisto, and Rhea, 1976-1979. Astron. J., vol. 85, July 1980, pp. 961-968.

2. Tüg, H., White, N. M., and Lockwood, G. W.: Absolute Energy Distributions of α Lyrae and 109 Virginis from 3295 Å to 9040 Å. Astron. Astrophys., vol. 61, 1977, pp. 679-684.

3. Jerzykiewicz, M., and Serkowski, K.: The Sun as a Variable Star. Lowell Observatory Bulletin, no. 137, 1966 (vol. VI, no. 18, pp. 295-323).

4. Lockwood, G. W.: Analysis of Photometric Variations of Uranus and Neptune Since 1953. Icarus, vol. 35, 1978, pp. 79-92.

5. Lockwood, G. W.: Secular Brightness Increases of Titan, Uranus, and Neptune, 1972-1976. Icarus, vol. 32, 1977, pp. 413-430.

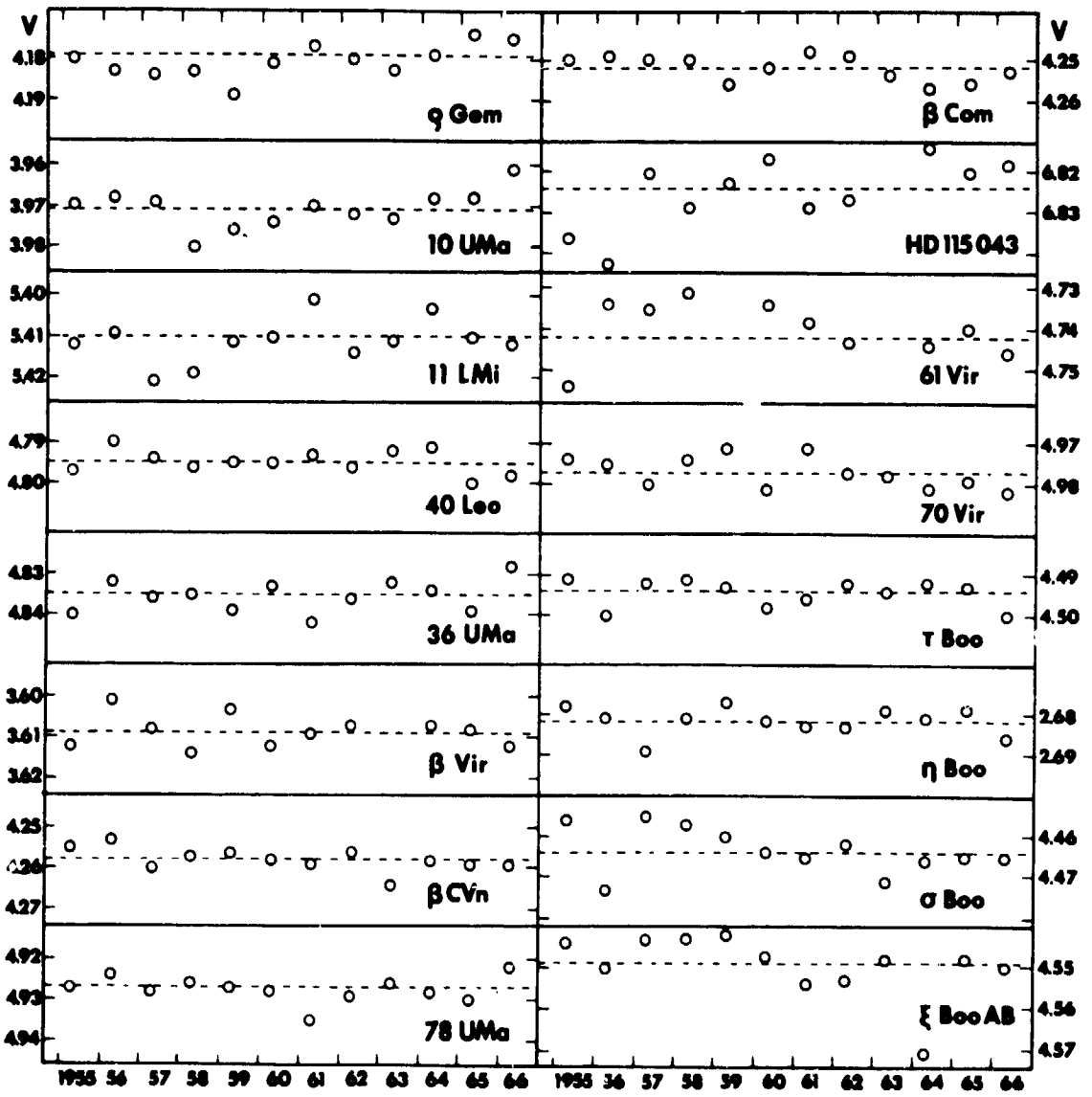


Figure 1. Annual mean V magnitudes of 16 F- and G-type dwarfs (from ref. 3).

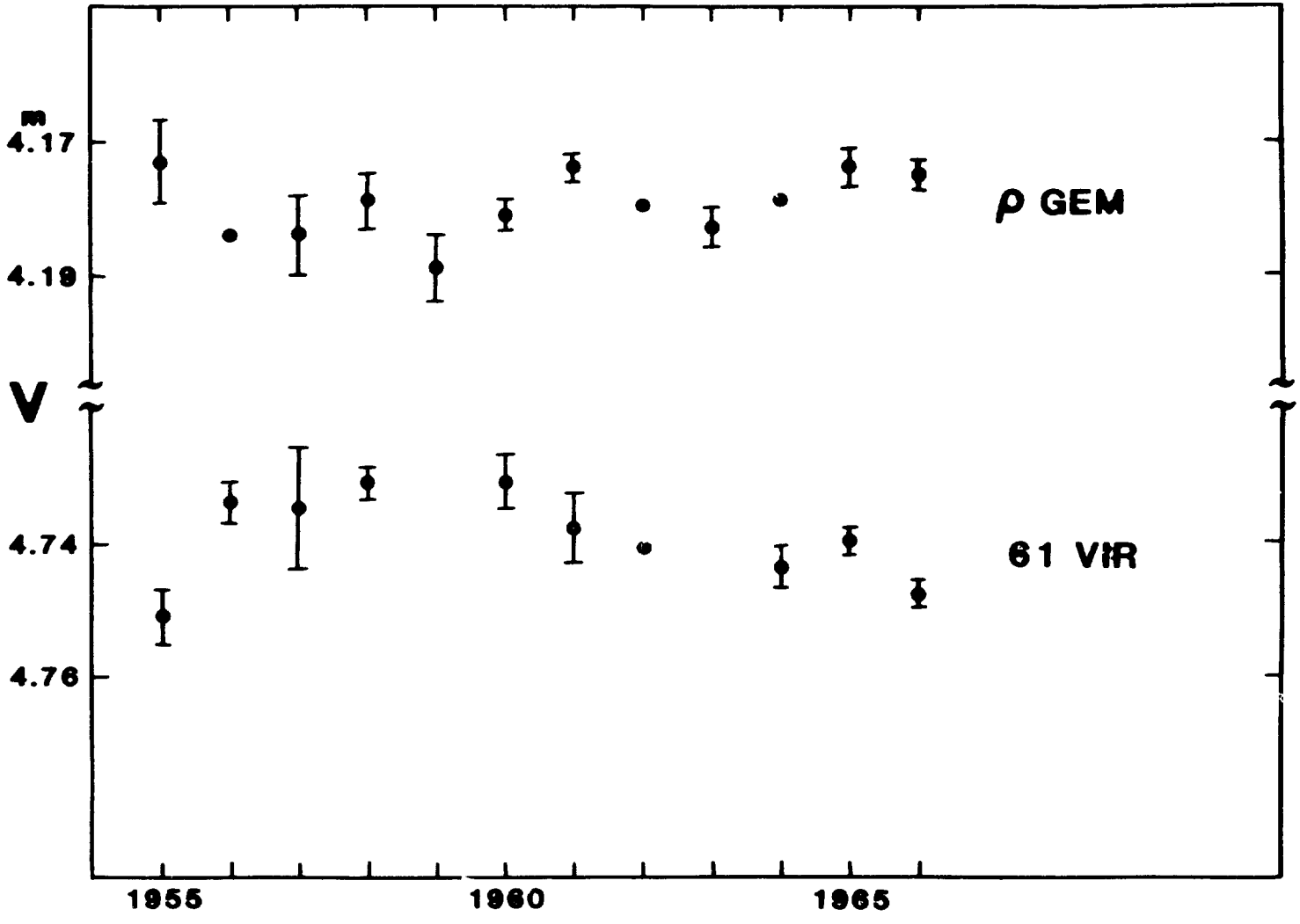
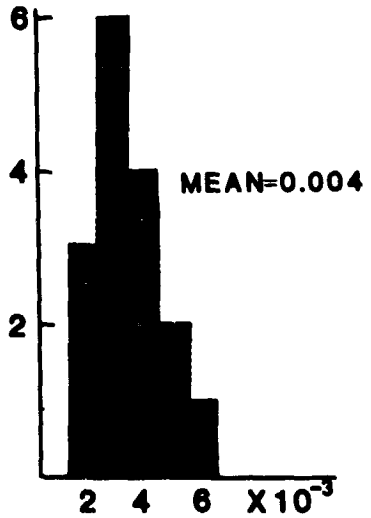
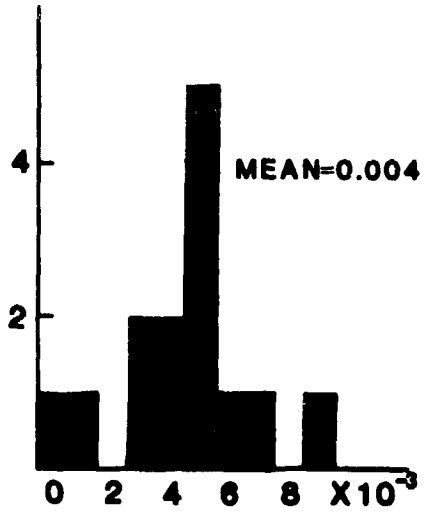


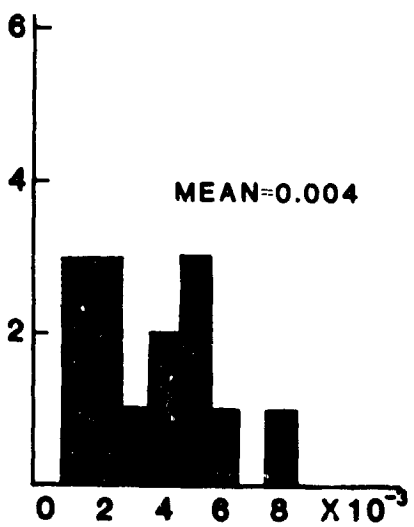
Figure 2. Annual mean V magnitudes of ρ Gem (FOV) and 61 Vir (G6V) showing the standard errors of the mean annual magnitudes. The average number of observations per year is 10.



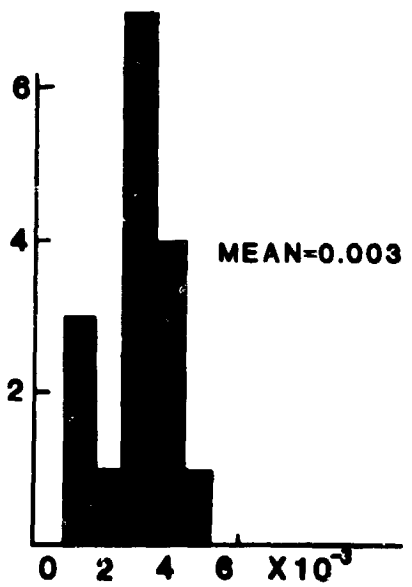
3a. Standard deviations of the annual mean V magnitudes of 16 F- and G-type dwarfs observed between 1962 and 1966



3b. Peak-to-peak ranges of the three-year mean y magnitudes of 14 F-, G-, and early K-type dwarf standard stars observed between 1972 and 1980



3c. Peak-to-peak ranges of the three-year mean y magnitudes of 14 standard stars, excluding solar-type dwarfs



3d. Standard deviations of the annual y magnitudes (reduced to the ensemble mean) of 16 F-, G-, and early K-type field stars observed between 1972 and 1980

Figure 3

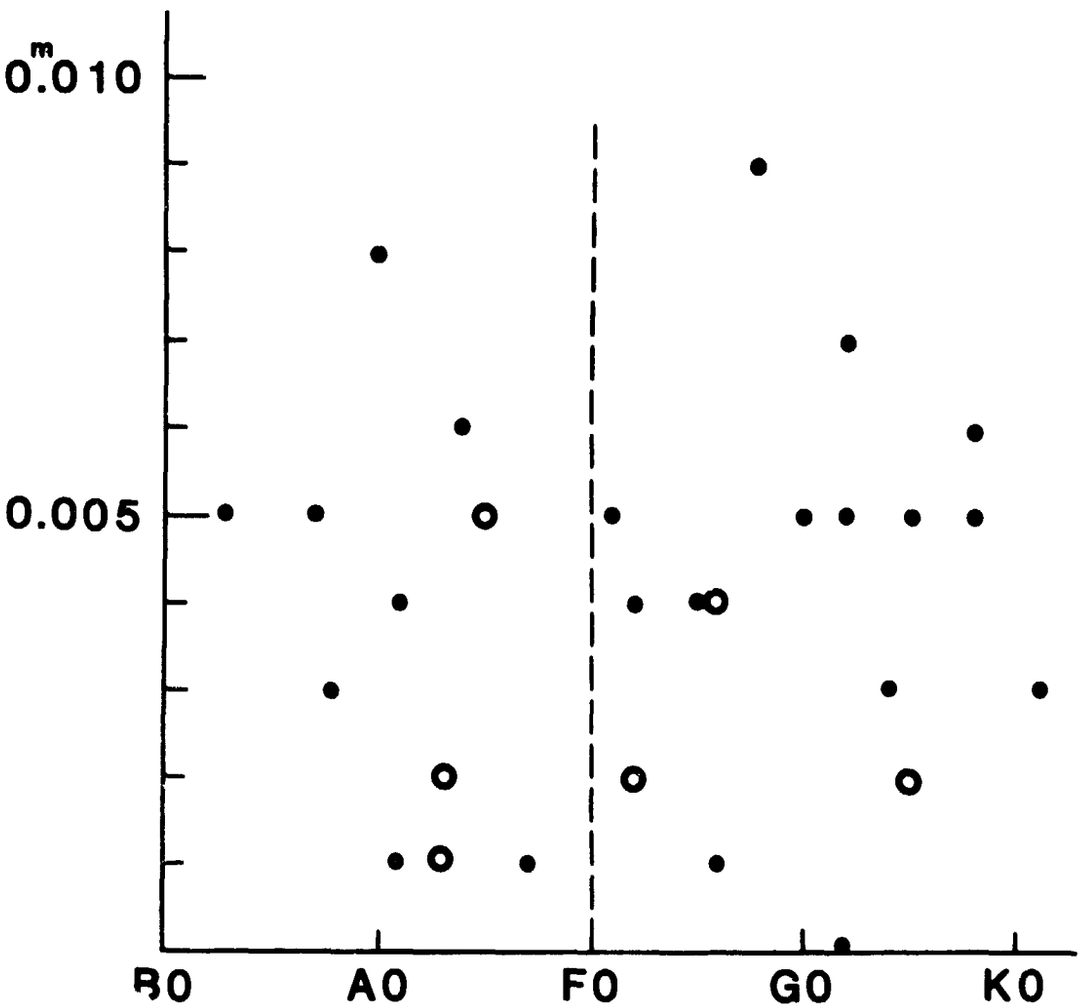


Figure 4. Peak-to-peak amplitudes of three-year mean y magnitudes of standard stars. (Open circles) - giants or unknown. (Filled circles) - dwarfs and sub-giants.

D22

PRECEDING PAGE BLANK NOT FILMED

N82 17036

SOLAR PULSATIONS AND LONG-TERM SOLAR VARIABILITY[1]

Philip R. Goode, Jerry D. Logan, and Henry A. Hill
Department of Physics
University of Arizona
Tucson, Arizona 85721

ABSTRACT

The seismology of the solar atmosphere is important in relating changes in luminosity to variations in other observables. This approach has already led to the identification of properties which were not previously observed or recognized. Equally important results from solar seismology are expected in the future.

INTRODUCTION

The relationship between variations in the sun's luminosity and long term variations in the earth's climate remains an open question in spite of considerable recent efforts to define it (see references 1 and 2 for a review). Various methods may be used to study the solar luminosity-terrestrial climate relationship. One approach to this study that is receiving serious attention is based on the observation of long-term variations of the solar diameter and related quantities. It is in establishing the relationship between these observables and solar luminosity where the study of solar pulsations can be important.

The fundamental measurements in a diameter-based program require a precise determination of the location of an edge of the solar disk. This location, as deduced from observation, depends on how the edge of the sun is defined. The edge obtained by any selected definition is sensitive, in general, to the shape of the solar limb darkening function. This function represents the decrease in the brightness of the sun as the limb is approached. The sensitivity of the edge location to time-varying whole-disk properties of this function was a factor in the misinterpretation of the Princeton solar oblateness measurements and later gave rise to equally serious questions concerning the reality of the solar oscillations reported at SCLERA[2]. The

---

[1]This work was supported in part by the National Science Foundation, the Air Force Office of Scientific Research, and the Department of Energy.

[2]SCLERA is an acronym for the Santa Catalina Laboratory for Experimental Relativity by Astrometry, a facility jointly operated by Wesleyan University and the University of Arizona.

228

existence of these oscillations has now been corroborated (reference 3) but in so doing, new questions have surfaced regarding stellar pulsation theory at a fundamental level. For discussions of this topic, see reference 4 for a review on solar oscillations and references 5, 6 and 7 for reviews on solar oblateness.

The limb darkening function reflects the temperature gradient in the solar envelope which in turn is evidenced by an outward energy transport, i.e., luminosity. It is clear that a change in the solar luminosity and, hence, in the temperature gradient implies an accompanying change in the limb darkening function. Because the limb darkening function can also be studied with considerable accuracy, an examination of its long-term variation would be expected to be as useful a diagnostic of the solar luminosity as an intrinsic diameter study alone. In any case, if solar diameter variations are to be used to investigate solar luminosity, then the link between changes in the limb darkening function and the observed diameter must be understood at a basic level.

The relationship between changes in the limb darkening function and luminosity variations is likely to be complicated. This possibility makes advisable the use of an empirical procedure in which changes in the limb darkening function are examined for correlations with direct radiometric measurements of the solar constant. This approach ought to be sensitive to correlations in changes in the solar constant of about one part in 10,000 with periods on the order of days. This accuracy is fixed by the current level of reproducibility of direct radiometric measurements. However, the extrapolation of the empirical calibration from a period of days to a period of years to decades is probably not justified. Accuracy of the extrapolation would be fortuitous since there is no compelling reason to believe that the physical processes leading to short-term changes are the same as those underlying long-term changes. This empirical calibration would be useful in providing comparative data for more analytic work to follow.

Efforts have been made to identify the primary sources of possible luminosity changes by modeling various convection zone effects. The work described in reference 8 exemplifies this approach. Although this work is important, it models the solar envelope in only the simplest manner, in spite of the fact that it is in the envelope that the complicated manifestations discussed here are observed. A different approach is pursued here -- an empirical one in which the effect on observed quantities of a change in the rate of energy transport through the solar atmosphere is studied without concern for the origin of the change. Studying changes in the limb darkening function with time and luminosity would be representative of this type of approach. Determining the empirical parameters of the observables and identifying the important physical processes has been and remains a major thrust of the theoretical and its associated observational work at SCLERA. Studies of this type are described in references 6, 9, 10, 11, 12, and 13.

Reference 12 addresses the nonlocal character of radiative transfer in the solar atmosphere. The work shows that this nonlocal character is of primary importance in the treatment of perturbations of variables describing the solar atmosphere. This formalism can be used to calculate the effect of a

disturbance like a gravity wave on the limb darkening function and the luminosity.

Work at SCLERA and the Sacramento Peak Observatory has resulted in a new observational technique to study the properties of pulsations in the solar envelope (reference 13). Use of this technique has led to the discovery of five-minute-period traveling waves in the photosphere which are probably gravity waves. This discovery is important because these waves may significantly affect the relationship between the solar limb darkening function and the luminosity. The technique utilized in this work determines the frequency and vertical, rather than horizontal, spatial characteristics of a disturbance by employing the observed Doppler shift of a spectral line as a function of position in the absorption line.

Intensity and displacement fluctuations arise in the essentially adiabatic interior of the sun, but are observed in the nonadiabatic atmosphere. In the photosphere, radiative damping should be the most important nonadiabatic process, directly affecting the properties of global oscillations in the photosphere where they are observed and, for instance, greatly affecting the propagation of traveling disturbances which may be responsible for heating throughout the atmosphere. It is therefore necessary to determine what types of disturbances occur in the photosphere and to be able to theoretically treat the radiative damping of such disturbances in detail.

RADIATIVE EFFECTS

The theoretical treatment of a disturbance in the atmosphere entails the study of the appropriate wave equation. Radiative damping enters the wave equation through the first law of thermodynamics as the heat gained by the system. Locally, this heat gain per unit time is given by the divergence of the radiative energy flux,

$$-\vec{\nabla} \cdot \vec{F}_R = \int_{4\pi} \kappa_\lambda \rho (J_\lambda - S_\lambda) d\lambda \quad ,$$

where $\vec{F}_R$ is the radiative flux, ρ the density, and κ_λ, J_λ, and S_λ the opacity, the mean intensity, and the source function, respectively, at wavelength λ. The opacity and source function are determined locally while the mean intensity is affected by other regions of the atmosphere. Reference 12 describes a method for addressing this problem and demonstrates the inadequacy of the standard (and local) Newtonian cooling law and the Eddington approximation. This inadequacy is manifest in their considerable failure to predict the radiative damping of perturbations in the photosphere, the place where such effects should be most important in the limb darkening function.

The method given in reference 12 shows that the nonlocal character of the mean intensity, in conjunction with the presence of line blanketing, leads to a wavelength dependence of the opacity that is also important in the calculation of radiative damping incorporating changes in the opacity. This nonlocal

approach yields more realistic spatial properties of disturbances in the solar atmosphere, allowing an improved comparison between the luminosity and limb darkening function variations. The calculation of these spatial properties is being pursued at SCLERA.

MECHANICAL EFFECTS

Mechanical energy flux has been observed in the photosphere (reference 13). The flux may be of consequence to the relationship between variations in the limb darkening function and luminosity.

The observations discussed here were performed at the Sacramento Peak Observatory. The data are high resolution line profiles of the 5434 Å Fe I line, which is a nonmagnetic line spanning the photosphere. By examining the Doppler shifts closer and closer to the bottom of the absorption line, the velocities of the disturbances are resolved at nine successively higher altitudes in the photosphere. These velocity data are filtered to examine power in the five-minute-period window.

Calculations using the filtered data reveal that there is a disturbance in the five-minute-period window in addition to the well-known five-minute-period acoustic mode. The secondary disturbance is found to be traveling waves which have ingoing and outgoing phase velocities. For both the ingoing and the outgoing traveling waves, power is observed for vertical wavelengths which are about 4/3 times the height of the atmosphere spanned by the observations. Longer wavelengths may be present but the current work has considerably reduced sensitivity in that region. Taking the height of the photosphere spanned by the observations to be 400 km, the vertical wavelength for both ingoing and outgoing phases is approximately 530 km. The period of these traveling waves is 278 ± 41 seconds and their vertical phase velocity is about 2 km/sec.

If the observed traveling waves comprise only a small portion of their spatial spectrum or if the waves are not localized, the group of waves which they represent may provide an important net vertical energy flux to the lower chromosphere. This flux, which may be as large as 10^7 ergs/cm^2/sec, could have implications for the relationship between the limb darkening function and solar luminosity. The horizontal extent of the traveling waves reported here is currently under investigation.

REFERENCES

1. White, O. R., ed. The Solar Output and Its Variation. University of Colorado Press, Boulder, Colorado, 1977.

2. Pepin, R. O., Eddy, J. A. and Merrill, R. B., eds. The Ancient Sun. Proceedings of conference held in Boulder, Colorado, October 1979.

3. Caudell, T. P., Knapp, J., Hill, H. A., and Logan, J. D. Nonradial and Nonlinear Stellar Pulsation, Lecture Notes in Physics No. 125, ed. H. A. Hill and W. A. Dziembowski. Springer-Verlag, Berlin, 1980, p. 206.

4. Hill, H. A. The New Solar Physics, ed. J. A. Eddy. Westview Press, Boulder, Colorado, 1978, Chapter 5.

5. Hill, H. A. and Stebbins, R. T. Astrophysical Journal, vol. 200, 1975a, p. 471.

6. Hill, H. A. and Stebbins, R. T. Fifth Cambridge Relativity Conference, Cambridge, Massachusetts, June 1974.

7. Hill, H. A., Stebbins, R. T. and Brown, T. M. Atomic Masses and Fundamental Constants, vol. 5, ed. J. H. Sanders and A. H. Wapstra. Plenum, New York, 1976, p. 622.

8. Dearborn, D. S. P. and Newman, M. J. Science, vol. 201, 1978, p. 150.

9. Hill, H. A., Rosenwald, R. D. and Caudell, T. P. Astrophysical Journal, vol. 225, 1978, p. 304.

10. Hill, H. A. and Caudell, T. P. Monthly Notices Roya' Astronomical Society, vol. 186, 1979, p. 327.

11. Knapp, J., Hill, H. A. and Caudell, T. P. Nonradial and Nonlinear Stellar Pulsation, Lecture Notes in Physics No. 125, ed. H. A. Hill and W. A. Dziembowski. Springer-Verlag, Berlin, 1980, p. 394.

12. Hill, H. A. and Logan, J. D. Astrophysical Journal, 1981 (submitted).

13. Hill, H. A., Goode, P. R. and Stebbins, R. T. 1981 (in preparation).

OBSERVATIONS OF LARGE-SCALE MOTIONS OF THE SUN

Barry LaBonte
Mount Wilson and Las Campanas Observatories
of the
Carnegie Institution of Washington

ABSTRACT

Recent observations of large-scale mass motions on the Sun are discussed. The principal large-scale velocity flows are convection, rotation, meridional flow, and torsional and radial oscillations.

INTRODUCTION

I would like to give a brief review of some recent observational results on the large scale motions of material on and in the Sun. This review is neither complete nor unbiased but is meant to remind people of the kinds of large-scale flows that may or may not be observable in the Sun. The optimistic assumption behind the study of large-scale motions is that large horizontal scales are in some way related to large vertical scales, i.e., great depths into the Sun. Thus, we hope that we are probing the inner working of the Sun. Stix (1981) has shown that large-scale velocity fields can be transmitted to the observable photosphere from great depths. However, if we see a large-scale motion, we cannot tell observationally at what depth it in fact originates. Such a determination requires interpretation in the context of a model. The Mount Wilson data I refer to have been taken by a series of observers under the guidance of Dr. Robert Howard, and with the support of the NSF, NASA, and ONR.

CONVECTION

In the solar convection zone there is radial outflow of hot gas and radial inflow of cool gas. The conversion of heat flux into gas motions is central to many solar phenomena.

The observable form of convection is the granulation. This is the only structure in which the radial velocity-temperature correlation is visible. Granule sized (about 2 arc seconds) are near the spatial resolution limit of velocity observations and therefore difficult to observe individually (Beckers and Morrison, 1970). If we are interested in the statistical properties of convection, however, we may observe a large-scale "velocity" pattern; the limb redshift. The velocity-temperature-intensity correlation in granulation causes the average wavelength of a spectrum line observed near the center of the solar disk to be blueshifted with respect to the wavelength observed near the limb (Beckers and Nelson, 1978). This effect amounts to several hundred meters per second if

interpreted as a Doppler shift, and is easily observable in low spatial resolution data (Figure 1).

By comparing the properties of the limbshifts observed in spectrum lines formed at various heights in the photosphere, some knowledge of the statistical character of granulation scale convection can be gleaned. From the comments of some of the other speakers it is clear that a more interesting question is whether the statistical character of convection varies with time. If so, then variations in the solar luminosity, radius, and neutrino flux might be expected. There are now 14 years of digital daily velocity maps of the Sun taken at Mount Wilson, and we are in the process of measuring the limbshift effect in that data to set limits on the magnitude of variation of granule convection.

Another velocity field which has been interpreted as an effect of convection is the supergranulation (Simon and Leighton, 1964). In supergranulation there are neither radial velocities (Giovanelli, 1980) nor temperature fluctuations (Worden, 1975), but only a cellular pattern of horizontal velocities. Supergranulation is of particular interest because magnetic fields on the solar surface are roughly organized into a network around the edges of supergranules. Supergranules are large enough (about 40 arc seconds) to be resolved in the Mount Wilson data, and the root-mean-square velocity amplitude in supergranulation is also being examined for time variations.

One other form of velocity pattern caused by convection has been predicted to be observable at the solar surface, namely, giant cells (Gilman, 1979). These cells would originate deep in the convection zone and have correspondingly large horizontal dimensions. We have searched the Mount Wilson velocity data for evidence of the existence of giant cells, but have not positively seen them, to limits of 3 to 10 m s^{-1} RMS amplitude (LaBonte et al., 1981). Model calculations predict amplitudes of this order, so better observations are of immediate importance. Some efforts to measure giant cell motions by using magnetic tracers have been made (Schöter and Wöhl, 1976), but the velocity sensitivity is not yet as good as the Doppler method.

ROTATION

The rotation of the Sun and the decrease of rotation rate with increasing latitude were known to the earliest telescopic observer from the motions of sunspots. The use of sunspots and other identifiable features as tracers of solar rotation remains a standard procedure. The equatorial rotation rate of recurrent isolated spots is 2.91±0.01μ rad s^{-1} (Newton and Nunn, 1951; Ward, 1966; Kearns, 1979; Clark et al., 1979; Neidig, 1980). Spot groups, which are younger on average than isolated spots, rotate 1 to 2% faster than isolated spots (Ward, 1966; Godoli and Mazzucconi, 1978; Kearns, 1979, Neidig, 1980; Wohl and Balthasar, 1980). There is no convincing evidence for systematic variations of spot rotation within a sunspot cycle or from one cycle to another, in the last 100 years;

reanalysis of spot drawings from the early 1600's suggests that variations might have occurred prior to the Maunder minimum (Eddy et al., 1977; Herr, 1978). Sunspots are excellent markers because of their large size and long lifetimes, but unfortunately exhibit peculiar motions which can reach 10 to 15% of the rotation velocity in extreme cases. The difference in rotation rates measured by single long-lived spots and by spot groups is a form of these peculiar motions. An additional problem is that spots are never seen at high latitudes. At Mount Wilson we have just begun a project to measure sunspot positions on the 70 years of direct photographs taken at the 60-foot tower.

A sunspot may be considered as a single unit of magnetic field. Unfortunately, photospheric magnetic fields outside spots are comprised of units, filigree, much too small and short-lived to measure individually. Larger groupings of filigree are long-lived, however, and rotation may be measured by using magnetic field patterns as tracers (Wilcox and Howard, 1970), We have recently looked at rotation in this way with Mount Wilson data and find the equatorial rate to be 2.91±0.01μ rad S^{-1} for each of the last 14 years. The magnetic field data may be used to measure rotation to the poles (Howard, 1978). A number of other magnetic tracers (white light faculae; chromospheric plages, and filaments; coronal emission patterns) have been used in the past, but the resulting accuracy has been low, since these features are ill-defined averages over many individually identifiable magnetic features.

There are only two tracers on the Sun which are not magnetic: granules and supergranules. Granules are too shortlived to be used, but Duvall (1980) has measured the equatorial rotation rate from the supergranulation velocity pattern to be 2.97μ rad s^{-1}, about the rate of young sunspot groups. This measurement should be repeated.

The other way to measure solar rotation is to observe the Doppler shifts of spectrum lines. In this type of data the rotation signal is by far the largest amplitude velocity pattern (Figure 1). The angular rotation rate can be measured at all latitudes, especially close to the poles. The observation of a strict monotonic decrease of rotation rate with latitude (Beckers, 1978a) sets limits on the structure of the convection zone (Gilman, 1979).

Because magnetic fields cover only a small fraction of the Sun, the Doppler measures essentially refer to nonmagnetic material. Doppler rotation measurements do not present a unified result. Stanford data (Scherrer and Wilcox, 1980a) give an equatorial rate of 2.90μ rad s^{-1}, the same as isolated recurrent sunspots. They also show no significant variation above the measurement noise on any timescale. Mount Wilson (LaBonte and Howard, 1981a) and Kitt Peak data (Livingston and Duvall, 1979; Duvall, 1981) give a rate 1 to 2% lower, and show variations on all timescales. Most of the day to day variations are caused by the instrument, but the origin of long term variations is not yet settled. The elimination of systematic errors is crucial to obtain an absolute rotation rate. There are differential measurements (Foukal, 1979) which suggest there is a

difference in rotation rate between magnetic and nonmagentic gas in the photosphere; more tests of this type should be done.

MERIDIONAL FLOW

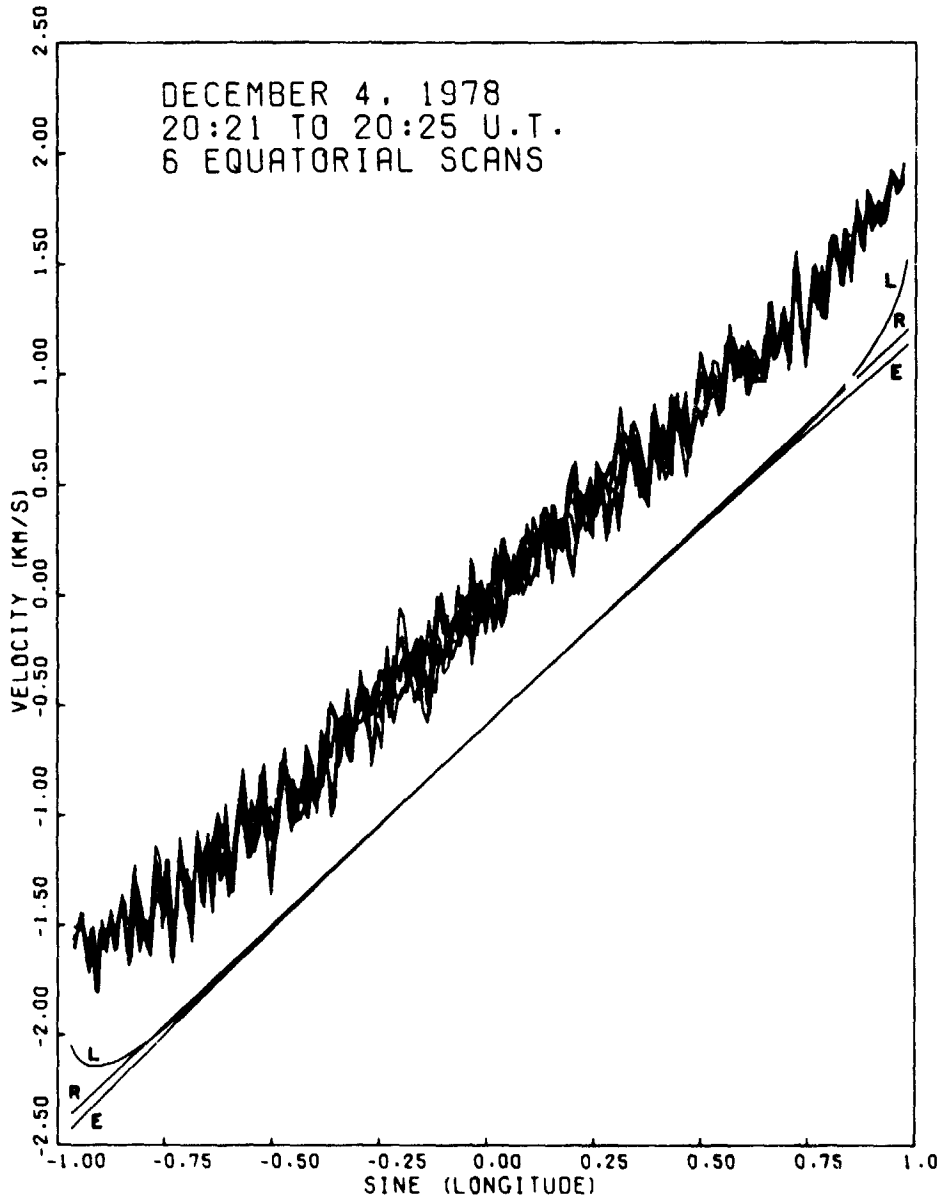
The differential character of the surface rotation implies that angular momentum is redistributed within the Sun, either radially, meridionally, or both. Theories differ on the amplitude and direction which is expected. The Doppler observations all agree that the flow is poleward, with amplitudes of $\simeq$ 20 m s^{-1} (Duvall, 1979) $\simeq$ 40 m s^{-1} (Beckers, 1978b) or $\simeq$ 15 m s^{-1} from Mount Wilson data. Beckers and Taylor (1980) have cautioned that meridional flow may be enhanced or masked by a latitude variation of the limb redshift (and thus, granulation), and their observations suggest that all the quoted values should be increased by $\simeq$ 30m s^{-1} poleward. Tracer measures of meridional flow are possible, using longlived magnetic patterns, and give a poleward flow $\simeq$ 10 m s^{-1} (Howard and LaBonte, 1981).

OSCILLATIONS

The role of short period oscillations in the Sun has been considered by other speakers, so I will restrict attention to long period oscillations. The sunspot cycle itself has long been thought of as a magnetic oscillation, but only recently has a velocity field been observed which shows some of the mass motions involved. This is the torsional velocity oscillation (Howard and LaBonte, 1980; Scherrer and Wilcox, 1980b; LaBonte and Howard, 1981b). The zones of magnetic (sunspot) appearance are found to be the shear zones of a torsional wave emanating from the poles toward the equator, with an 11 year period and 22 year travel time. Thus 2 waves are visible on the Sun at all times. There is also a lower wavenumber torsional oscillation with a period of 11 years, which has the appearance of a periodic steepening and flattening of the differential rotation curve (Livingston and Duvall, 1979; LaBonte and Howard, 1981b). These wave modes have amplitudes 3 to 10 m s^{-1}. It is possible that other torsional modes exist but are of lower amplitude. No meridional component of these waves is detected with limits $\simeq$ 30 m s^{-1}.

As other speakers have indicated, it is not known whether the solar radius varies. Low amplitude radial oscillations are not ruled out by existing data. At Mount Wilson we are testing this question in two ways. From our daily full disk magnetic observations we obtain objective radius values, with a precision of $\simeq$ 0.25 arc seconds. We are now analyzing the past 8 years of data to search for radius changes. Second, the project to measure the 70 year series of direct photogaphs will include measures of the radius. Tests indicate that a precision <1 arc second per plate should be possible, and systematic errors uniform and controllable for the entire dataset. The plate measuring project will take about 2 years.

Beckers, J.M.: 1978a, Astrophys. J., 224, L143.
Beckers, J.M.: 1978b, in G. Belvedere and L. Paterno (eds.), "Workshop on Solar Rotation", Catania, p. 166.
Beckers, J.M. and Morrison, R.A.: 1979, Solar Phys., 14, 280.
Beckers, J.M. and Nelson, G.D.: 1978, Solar Phys., 58, 243.
Beckers, J.M. and Taylor, W.R.: 1980, Solar Phys, 68, 41.
Clark, D.H., Yallop, B.D., Richard, S., Emerson, B. and Rudd, P.J.: 1979, Nature, 280, 299.
Duvall, T.L.: 1979, Solar Phys., 63, 3.
Duvall, T.L.: 1980, Solar Phys., 66, 213.
Duvall, T.L.: 1981, B.A.A.S., 12, 896.
Eddy, J., Gilman, P.A. and Trotter, D.E.: 1977, Science, 198, 824.
Foukal, P.: 1979, Astrophys. J., 234, 716.
Godoli, C. and Mazzuccioni, F.: 1978, in G. Belvedere and L. Paterno (eds.), "Workshop on Solar Rotation", Catania, p. 135.
Gilman, P.A.: 1979, Astrophys. J., 231, 284.
Giovanelli, R.G.: 1980, Solar Phys., 67, 211.
Herr, R.B.: 1978, Science, 202, 1079.
Howard, R.: 1978, Solar Phys., 59, 243.
Howard, R. and LaBonte, B.J.: 1981, Solar Phys., in press.
Howard, R., Boyden, J.E. and LaBonte, B.J.: 1980, Solar Phys., 66, 167.
Kearns, M.D.: 1979, Solar Phys., 62, 393.
LaBonte, B.J. and Howard, R.: 1981a Solar Phys., in press.
LaBonte, B.J. and Howard, R.: 1981b Submitted to Solar Phys.
LaBonte, B.J., Howard, R. and Gilman, P.A.: 1981, submitted to Astrophys. J.
Livingston, W. and Duvall, T.L.: 1979, Solar Phys., 61, 219.
Neidig, D.F.: 1980, Solar Phys., 66, 205.
Newton, H.W. and Nunn, M.L: 1951, M.N.R.A.S., III, 413.
Scherrer, P.H. and Wilcox, J.M.: 1980a, Astrophys. J., 239, L89.
Scherrer, P.H. and Wilcox, J.M.: 1980b, B.A.A.S., 12, 473.
Schröter, E.H. and Wöhl, H.: 1976, Solar Phys. 49, 19.
Simon, G.W. and Leighton, R.B.: 1964, Astrophys. J., 140, 1120.
Stix, M.: 1981 Astron. Astrophys., 93, 339.
Ward, F.: 1966, Astrophys. J., 145, 416.
Wilcox, J.M. and Howard, R.: 1970, Solar Phys., 13, 251.
Wöhl, H. and Balthasar, H.: 1980 Astron. Astrophys., 92, 111.
Worden, S.P.: 1975, Solar Phys., 45, 521.



1) Raw Doppler velocity data. Measured line of sight velocities from 6 scans along the solar equator are plotted. The total time interval was 4 minutes. The smooth curves (displaced downward by 0.5 km s^{-1}) labelled R, E, and L are respectively the solar rotation; rotation plus "ears" (Howard et al., 1980); and rotation plus ears plus limbshift, as measured from the data. The velocity zero is arbitrary; positive velocities are motions away from the observer. Spatial resolution is 12.5 arc seconds, or 0.013 in sine longitude. Five minute oscillations dominate the velocity variations for |sine longitude |<0.5, and supergranulation for | sine longitude |>0.5.

THE COMBINED SOLAR AND TIDAL INFLUENCE IN CLIMATE

P. R. Bell
Institute for Energy Analysis
Oak Ridge Associated Universities

ABSTRACT

To provide an early warning indication of the CO_2 warming signal, we are searching for periodic or projectable trends in climate. The strong 20.5 year oscillation in Eastern North American January temperature found by Mock and Hibler shows evidence of a beat between waves with periods of 22.36 (22.21 to 22.55) years and 18.64 (18.45 to 18.79) years with an opposition at about 1880. These are interpreted to be the 22.279 year solar Hale magnetic cycle and the 18.61 year lunar nodal tidal cycle. The lunar nodal cycle is known to produce changes in the sea surface temperature through increased mixing of the mixed layer of the ocean. This beat note is shown to be evident in the Western High Plains drought record of Mitchell, Stockton and Meko and to provide a better fit to the drought series, especially at the beat oppositions in 1880 and 1770.

INTRODUCTION

An important part of the research into the effects of increasing CO_2 in the atmosphere is the early detection of the effect of CO_2 upon climate. Detection is important because any change in man's production of CO_2 will be slow to implement and because the response to change has a delay of perhaps a decade before it can become effective. This is a difficult problem because of the considerable natural variation of the climate with time scales of a decade or more.

It is clear that if the rhythmic and long-term trends in atmospheric temperature are ignored, the influence of CO_2 will go undetected until the year 1990 or 2000. See for example Madden and Ramanathan (1), who made such a study ignoring any trends in temperature.

Several authors have chosen a more fruitful course and have attempted to make predictions of current and future climate based upon the long temperature records with strong regular variations available from ice cores from the Greenland ice cap. See for example Dansgaard et. al. (2), Hibler and Langway (3), and Broecker (4).

Figure 1 is a reproduction from Broecker's paper and shows the temperature projection based on the Camp Century cycles. The climate projections of Dansgaard et al. and Hibler and Langway are similar. All of the projections based on ice core data illustrate the nature of the difficulty involved in detecting the CO_2 effect, as they all suggest that the temperature in the 1980s and 1990s should be low, rather like that of 1810-35 which was perhaps 0.5°C below the mean temperature of the last 200 years. The change is a result of natural processes. This unfortunate occurrence of an expected extended cool period just when the CO_2 effect should become apparent emphasizes the need to understand, if possible, any and all predictable regularities in the climate. Even if these components are not large enough to be dominant, they could reduce the range of uncertainty in the expected temperature projection and thus allow a more certain detection of the CO_2 effect.

THE BEAT EFFECT IN ATMOSPHERIC TEMPERATURE

It has long been known that at least Eastern U. S. winter temperatures showed a considerable regular variation (5). Recent analysis by Mock and Hibler (6) of January temperatures over eastern U. S. and Canada shows a strong regular oscillation in these temperatures with a mean period of 20.5 years. There is also remarkable coherence among the separate records from the 12 stations analyzed. Figure 2 is a reproduction of Figure 3a of their paper (7).

Mock and Hibler's plot has the classic form of a beat note between two regular oscillations of comparable frequencies. If we assume that a beat is involved, the two oscillations were clearly in maximum reinforcement in about 1935-40 and were in opposition in about 1880. The oscillations are approaching another opposition in the immediate future. About three peaks fall between the maximum coherence (and amplitude) and the opposition giving about 6 waves to the full beat. The actual number need not be an integer, but is obviously in the range $6 \pm .5$. From the characteristic shape of the opposition interval around 1880 it is clear that the shorter period (higher frequency) wave is somewhat stronger than the longer period wave and the longer period wave must have exactly one fewer cycles between oppositions. The long period wave thus completes $5 \pm .5$ oscillations while the short period wave completes $6 \pm .5$. Given the mean period of 20.5 years from Mock and Hibler's analysis, it can be shown algebraically that the two oscillations involved have periods of 22.36 (22.21 to 22.55) and 18.64 (18.45 to 18.79) years (8).

These periods are remarkably close to the solar magnetic oscillation cycle of $22.279 \pm .027$ years established by Dicke (9) and the lunar nodal cycle of 18.61 years (established astronomically with negligible error). Both cycles are plausibly involved with observed climate, the phase corrected solar magnetic sunspot period via a solar luminosity change with a peak to peak amplitude of 0.3% or less, [Dicke (9)], and the lunar nodal cycle (the period of rotation of the plane of

the moon's orbit) through the modulation of the twice daily ocean tidal currents as explained by Loder and Garrett (10). Marine effects of the 18.6 year period lunar phenomena have been reported by Hachey and McLellan (11) and Maximov and Smirnov (12) for sea surface temperatures and by Maximov and Sleptsov-Shevlevich (13) on arctic sea ice area. Indirect evidence of changes in the temperature of the sea with the 18.6 year period are shown by the shifts in the latitude of the southern limit of wintering of herring in the North Atlantic in phase with the lunar declination cycle from Kislyakov (14). The catch of striped bass off the U.S. east coast also shows opposite phase variations at northern and southern ports indicating migration of fish populations in phase with the lunar declination [Rust and Kirk (15)]. With such marked and ubiquitous effects in the marine environment, it would be surprising if the 18.6 year lunar cycle did not show up in atmospheric records as well.

THE BEAT EFFECT IN WESTERN HIGH PLAINS DROUGHT

The postulated interaction between solar magnetic and lunar nodal periods also allows a more satisfactory analysis of droughts on the U.S. Western High Plains. The careful work of Mitchell, Stockton, and Meko (16) demonstrates a strong and predictable cycle in such droughts. They show an interesting correlation between drought and solar magnetic cycles. In addition, they demonstrate a rather good correlation between the envelope of sunspot numbers (a general indicator of long term solar activity) and the intensity of drought as indicated by their Drought Area Index (DAI). Droughts are more severe when the solar activity is highest.

Figure 3, adapted from Mitchell, Stockton, and Meko's figure 2, displays drought indices derived from tree ring data for the High Plains area. I have added a set of regular time marks at the spacing of the solar magnetic cycle, 22.279 years. Note that the marks and peaks match rather well except for the time around 1800 where the waves shorten and allow the insertion of an extra wave. This, again, is the typical effect of a beat where the shorter period is more intense than the longer. Thus, the three droughts 1862, 1882 and 1900-01 do not correlate well with the solar cycle marks, the 1882 drought being well between two marks and also of smaller intensity. (The time around 1880, recall, is also the time when opposition of the solar and lunar waves is evident in Figure 2.) There are 12 full waves lying between the drought peak at 1711.4 and that at 1955.7 giving a mean drought period of 20.36 years (17). Inclusion of the latest drought at about 1976.5 changes this but little to 20.39 years (18).

An elementary algebraic analysis similar to that performed for the Eastern January air temperatures yields the periods of the two component waves. Again, there are six waves of the shorter period oscillation and five waves of the longer oscillation between the oppositions at about 1880 and 1770. Together with the gross period of 20.36 years from the

figure of Mitchell, et al. (16), this yields 22.21 and 18.51 years for the two components of the drought cycles. This is in good agreement with the estimates from the January air temperatures, and the solar and lunar cycles.

The evidence reviewed here suggests the hypothesis that beat interactions between effects of the solar magnetic and lunar nodal cycles may account for a significant proportion of observed climatic variability on the decadal scale.

This hypothesis can be further explored through a simple model in which the lunar nodal cycle and the solar magnetic cycle are allowed to interact to create a single beat cycle. In the version described here, the amplitude of the lunar cycle is assumed to be constant, while the amplitude of the solar cycle is adjusted each year by the historically recorded sunspot number envelope (19), reflecting the findings of Mitchell, et al. (16). The model then has four remaining parameters: the period of each cycle, their relative amplitudes, and their relative phase. The period parameters are fixed by hard astrophysical data: the 18.61 year lunar tidal period and the 22.279 year solar magnetic period. The observed January temperature records cited earlier fix the phases in opposition at about 1880. The relative amplitude cannot be specified independently, because we have not suggested which causal interactions are involved in the postulated relationship. Instead, we adjust the amplitude of the lunar wave empirically with respect to the solar wave to provide an opposition effect similar in pattern to that shown in Figure 2 for January temperatures. This yields an amplitude for the solar wave about 2/3 that of the lunar wave at the 1880 opposition.

The beat wave resulting from this model is shown in Figure 4. The peak of each wave is marked by a vertical line. Just below these are plotted the times of maximum drought intensity from Mitchell, Stockton and Meko as shorter vertical lines. Note the rather good agreement with the peak times, including that for the extra wave that arises at the 1880 opposition. There is another opposition at about 1770 at which time the solar influence, adjusted to reflect sunspot records, was stronger than the lunar tidal curve and no extra peak was produced. In the nodal pattern, however, a wider peak spacing appears. This fits the wider spacing of the drought data maxima at 1757 and 1781.

A consequence of this beat scale is that the coming opposition that should fall near 1991 should in all probability be like that near 1770 (that is, without an extra drought inserted) since the general intensity of solar activity, as indicated in the sunspot envelope, is now much higher than it was in 1880. Thus the next High Plains drought is predicted to occur in about AD 2005, as shown in the figure. If any drought is noticeable in the interval it should be very weak and at about 1991. It must be recognized, however, that the increasing CO_2 heating may affect North American precipitation to such an extent that the High Plains droughts may no longer be recognizable among other more widespread drought conditions in these latitudes.

The beat wave of Figure 4 is rather different in shape from the filter output of Mitchell, Stockton and Meko seen in Figure 3. This is because the beat wave is entirely linear, being the sum of two sine waves, while the drought area index is highly non-linear, that is an area is either above or below the condition for drought state-1. When drought increases in an area, the count for the index may only increase if the drought area expands. This gives a considerable limitation to the amplitude range of the DAI.

Mitchell, Stockton and Meko also used a harmonic dial diagram to illustrate the phase correlation between the Hale magnetic sunspot cycle and the droughts. Figure 5(a) illustrates their figure. Here the time between one Hale sunspot minimum and the next (about 20-24 years) is spread uniformly around one turn of the dial and the droughts falling in this interval are plotted at an angle appropriate to their time of occurrence in this time interval. The distance from the center is proportional to the drought intensity insofar as it is indicated by the filtered Drought Area Index. The three droughts 1862, 1882, and 1900-01 can be seen to have large phase errors compared with the rest. A similar harmonic dial was prepared (Figure 5[b]) using the peak times of the beat wave (Figure 4) as the timing intervals for the dial's rotation. Here only the means of the drought peaks from the two separate filters of Mitchell, Stockton and Meko are plotted for simplicity. The plot now has no large phase errors. There is one and only one drought peak for each rotation of the dial. The purpose of this illustration is to display the improved phase distribution resulting from the beat wave timing.

These fairly strong indications of the involvement of the 18.61 year tidal modulation in the climate suggest the advisibility of looking carefully for this period in the recently available Pacific sea surface temperature series and in any other available long time climate records. Several such records are suggested: 1) The D/H ratio of bound hydrogen in tree ring cellulose already known to show the 22 year solar cycle [Epstein and Yapp (20)]; 2) the shorter periods in the Greenland ice core records found by Hibler and Johnson (18) would benefit greatly from a filter analysis; 3) Atlantic sea surface temperature records (such as are available). The particular component of the lunar tides that is strong in the Atlantic is not the same as that of the Pacific (semi-diurnal for Atlantic and diurnal for the Pacific) [Loder and Garrett (10)]. In theory, these components should be modulated with opposite phases of the 18.61 year cycle; 4) European or North African winter temperature records, since these are affected mostly by Atlantic sea surface temperatures, whereas U.S. temperatures are affected more by Pacific surface temperatures; 5) Total U. S. water supply as reflected in mean stream flow which has recently been shown to be strongly periodic, [Langbein and Slack (21)]; and 6) A similar study of European or African total water supply.

TIDAL EFFECTS IN PRECIPITATION

If the above considerations suggest strongly that there is a tidal influence in climate, then it would be surprising if there were not an influence of the stronger 14 day cycle of the tides, even though it is briefer and has less time for its effects to accumulate. A search of the literature located a group of papers from 1962 to 1969 [Bradley, Woodbury and Brier (22); Adderly and Bowen (23); Brier and Simpson (24) and references therein] on tidal effects. These papers document a correlation of heavy precipitation with the lunar synodic (phase) month, the anomalistic (perigee) month and the tropical (declinational) month. These are supported by records from 1500 U.S. weather stations for 50 years and from 50 New Zealand stations for 24 years. These reports reinforce the conclusion that tidal currents affect sea surface temperature and hence precipitation. Brier and others have suggested that these effects may be due to the atmospheric tides. It seems more likely to me that the effects are due to the ocean tides as the oceans have much greater energy storage while the atmospheric tides are very small.

CONCLUSION

The interacting influence of a solar magnetic period of 22.3 years and a lunar tidal period of 18.6 years (as well as shorter periods) in climate is strongly suggested. These influences are best seen not as global effects but in large area winter temperatures and in droughts in large areas. From the nature of the increased tidal mixing effect on the ocean mixed layer, it would be expected that times of high tidal motion would produce colder summer sea surface temperatures, while periods of lower tidal motion would give warmer summer sea surface temperatures. The opposite phase of the effect in the Atlantic and Pacific should provide strong regional effects, but latitude zonal averaging should greatly obscure the overall effect in both temperature and precipitation. Zonal averaging will thus generally conceal these rather prominent effects. The beat between the lunar and solar influences will produce periodic oppositions at about 111 year intervals so that the oscillations will not be constant in amplitude. The effects of the opposite phase of the tidal modulation in Atlantic and Pacific should cause the oppositions to occur alternately in the climate in regions dominated by each ocean. Several sources of surface temperature and climatic records need to be examined further for evidence of the combined solar and tidal effects.

The 18.6 year, the 22.3 year and the Camp Century ice core periods were combined in appropriate proportions, producing a time series that does confirm the principal conclusions of the Broecker paper (Broecker, 1975 [4]), that the CO_2 warming effect is not yet seen because the other driving forces of the climate are producing a tendency to a cold interval at present. The intermediate and longer cycles in the climate need better confirmation, however, before a firm conclusion can be established.

REFERENCES

1. R. A. Madden and V. Ramanathan, Science, 209, 763-769 (1980).

2. W. Dansgaard, S. J. Johnsen, H. B. Clausen, and C. C. Langway, Jr., in The Late Cenozoic Glacial Ages, K. K. Turekian, Ed. (Yale University Press, 1971), pp. 38-58.

3. W. D. Hibler, III, and C. C. Langway, Jr., SCOR/SCAR Polar Oceans Conference Montreal (1974).

4. W. S. Broecker, Science, 189, 466 (1975).

5. J. Spar and J. S. Mayer, Weatherwise, pp. 128-130, (June, 1973).

6. S. J. Mock and W. D. Hibler, III, Nature, 261, 484 (1976).

7. This figure deserves some comment. Filters can produce, on their own, some complex output shapes, but only when the filter output is very small compared with the input signal. This is not the case here. A single time series result is significant but here there are 12 separate time series from independent stations across eastern North America. All the separate outputs are essentially identical, an event hardly likely to occur by chance. The beat opposition seen in the figure occurs at the same time as that seen in the western U.S. drought series analyzed below. The probability of the coincidence is also low.

8. Define 'a' as the number of oscillations between oppositions in the short period wave. 'a-1' is then the number of oscillations between oppositions in the long period wave. 'p' is the mean period, 'p_1' the period of the long wave component, and 'p_2' the period of the short wave component. Then since $(a-1)*p_1 = a*p_2$ and $(p_1 + p_2)/2 = p$, it follows that $p_1 = [a/(a-0.5)]*p$ while $p_2 = [(a-1)/(a-0.5)]*p$. The quoted periods and their asymmetrical "confidence limits" result from using values of a = 6 - .5, and 6+.5 in these equations.

9. R. H. Dicke, Nature, 280, 24 (1979).

10. The effect of the twice monthly variation of the tides is prominent in producing changes in sea surface temperature in shallow waters, and in many estuaries the water column goes from strong stratification at the neap tide to essentially complete uniformity during the spring tides each month. The effect of the 14 day tide variation on coastal water temperature is quite strong. (R. H. Herlineaux, Fish. Res. Bd. Can. Progr. Rept. Pac. Coast Sta., 108, pp. 7-9 [1957].) The change in the moon's position over the nodal cycle ($\pm$ 18° to $\pm$ 28° in the sky) has a modulation effect upon tidal currents

of 10 to 15% (J. W. Loder and C. Garrett, J.G.R., 83, No. C4, p. 1967 [1978] and G. Godin, The Analysis of Tides [University of Toronto Press, 1972]). As explained by Loder and Garrett, the tidal currents move massive quantities of coastal waters over irregular bottoms twice a day, producing variable eddy diffusion which modulates the degree of mixing of the ocean mixed layer and hence the sea surface temperature.

11. H. B. Hachey and H. J. McLellan, J. Fish. Res. Bd. Can., 7, p. 355, (1948). The results of Hachey and McLellan have been verified and extended by an examination of the Bay of Fundy temperature records for 1921 to 1980. An 18.61 year period with an amplitude of 0.9°C is prominent in the water temperature record.

12. I. V. Maximov and N. P. Smirnov, Acad. of Sci. of the USSR: Oceanology, 5, No. 2, pp. 15-24 (1965).

13. I. V. Maximov and B. A. Sleptsov-Shevlevich, Tr. Polyar Nauch. Issled. Inst. Mosk. Ryb. Khoz. Okeanogr. (Pinro), 27, pp. 21-39 (1970).

14. A. G. Kislyakov, reported by Maximov and Smirnov, Ref. 11.

15. B. W. Rust and B. L. Kirk, in Time Series and Ecological Processes, Proceedings of a SIMS Conference, June 27-July 1, 1977.

16. J. Murray Mitchell, Jr., C. W. Stockton, and D. M. Meko, in Solar-Terrestrial Influences on Weather and Climate, B. M. McCormac and T. A. Seliga, Eds. (Reidel Pub. Co., Holland, 1979), pp. 125-143.

17. Despite a gross period of 20.36 years, Mitchell, et al.'s Fourier analysis of the DAI series gave a period of 22 years. The 22 year component over most of the time interval is stronger than the 18.6 year period and the amplitude limiting effect of the non-linearity of the DAI enhances the ability of the Fourier transform to ignore short interferences with a moderately regular series.

18. It is interesting to note that a predominant 20 year period was found in the oxygen isotope record for 3 Greenland ice cores over a 728 year period by W. D. Hibler, III, and S.J. Johnson (Nature, 280, 481-3 [1979]). The oxygen isotope variations are related to the atmospheric temperature at the time of snowfall.

19. Following Dicke (9) the sunspot number amplitudes were increased by a factor of 1.3 from the beginning of the record until 1816 after which it was decreased linearly to 1.0 at 1837. If this correction is omitted or made as large as 2.0 there is no significant change in the results. This correction allows for the undercount in early observations due to the use of small telescopes. See K. O. Kiepenheur, The Sun (Ed. G. R. Kuiper), University of Chicago Press, 1953.

20. S. Epstein and C. J. Yapp, Earth and Plan. Sci. Ltrs., 30, 252-261 (1976).

21. W. B. Langbein and J. R. Slack, USGS Water Resources Rev. January (1980).

22. D. A. Bradley, M. A. Woodbury and G. W. Brier, Science, 137, pp. 748-749 (1962).

23. E. E. Adderley and E. G. Bowen, Science, 137, pp. 749-750 (1962).

24. G. W. Brier and J. Simpson, Quar. J. Royal Meteor. Soc., 95, pp. 120-147 (1969).

I gratefully acknowledge the help of Dr. W. C. Clark with revisions of this paper.

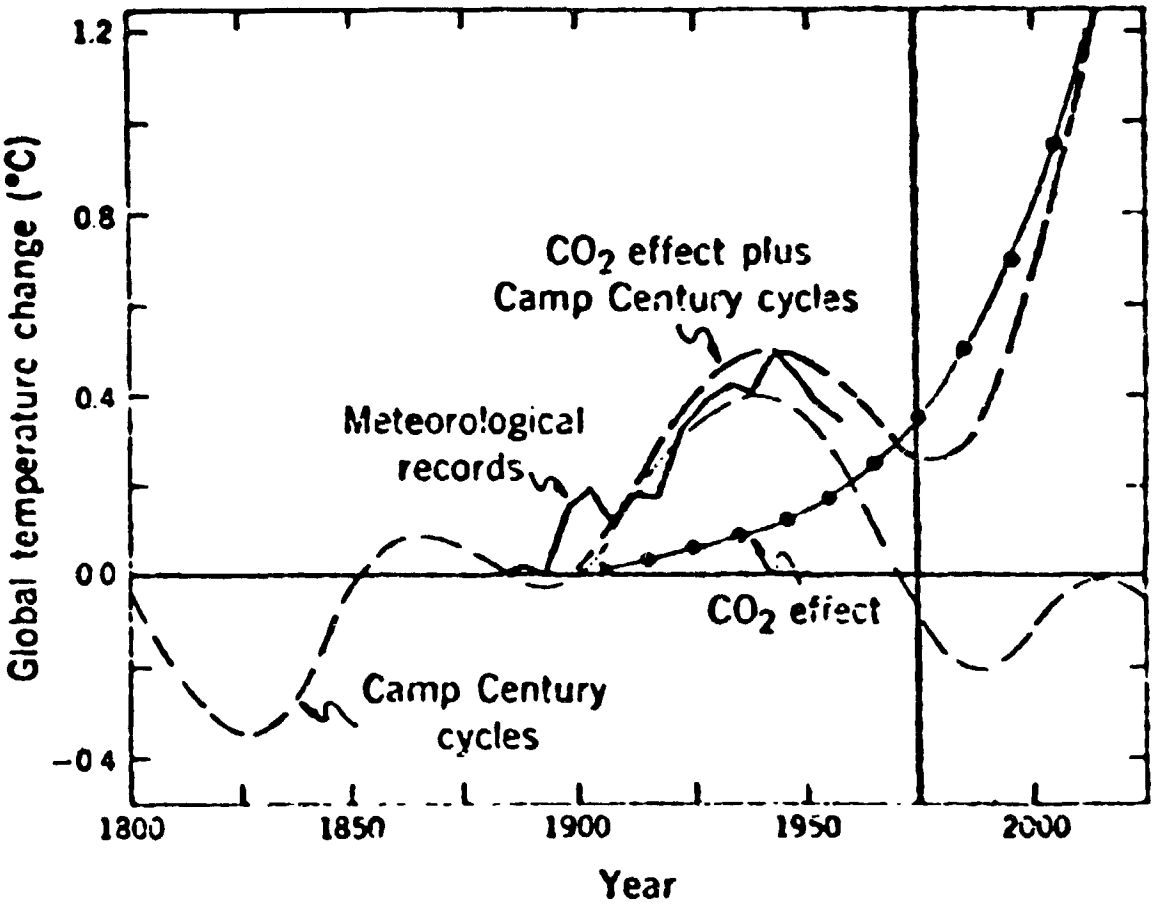


Figure 1. Broecker's temperature prediction based on the Camp Century ice core. The lower dashed curve is the ice core prediction alone. The upper dashed curve is the same prediction with the indicated CO_2 effect added.

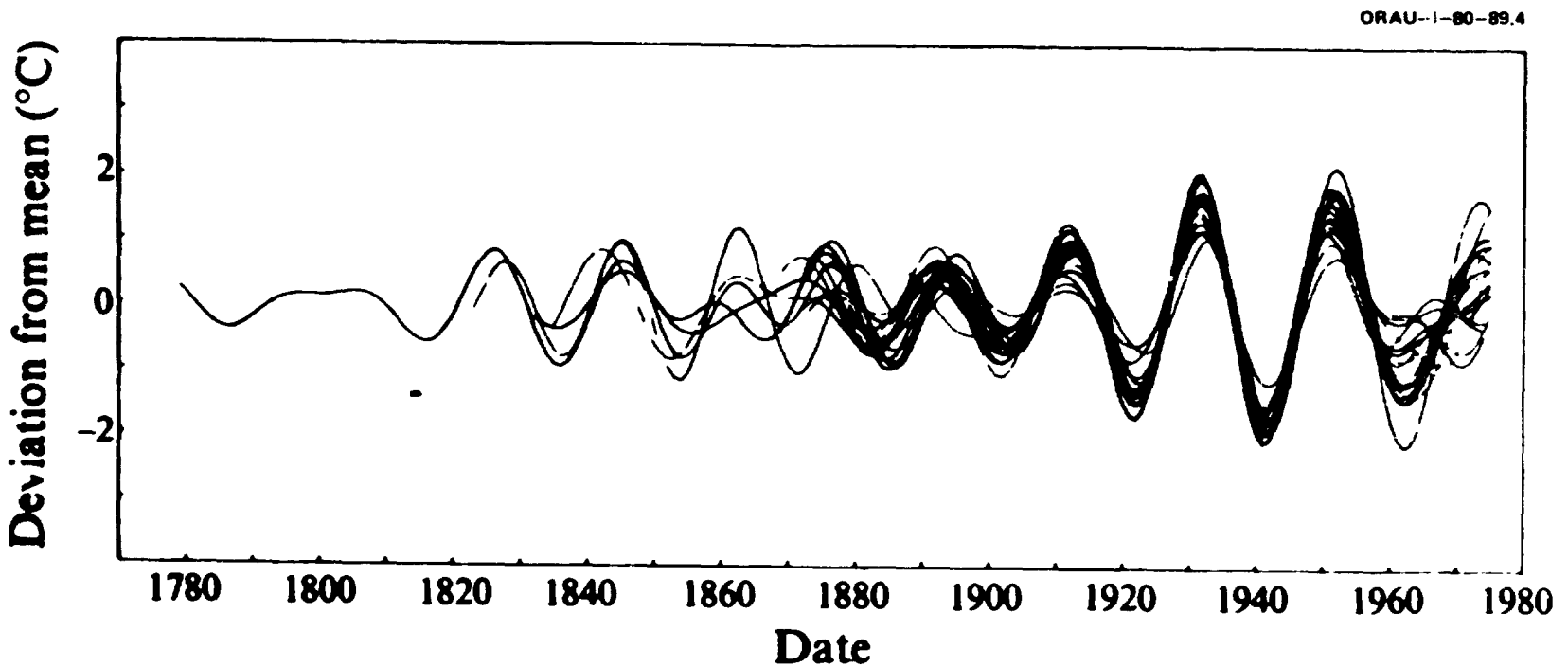


Figure 2. Mock and Hibler's superposition of 12 Eastern U.S. temperature records for January filtered through a band pass filter around 20.5 years. The separate records show remarkable coherence and strongly resemble the beat between two waves with comparable amplitudes and frequencies.

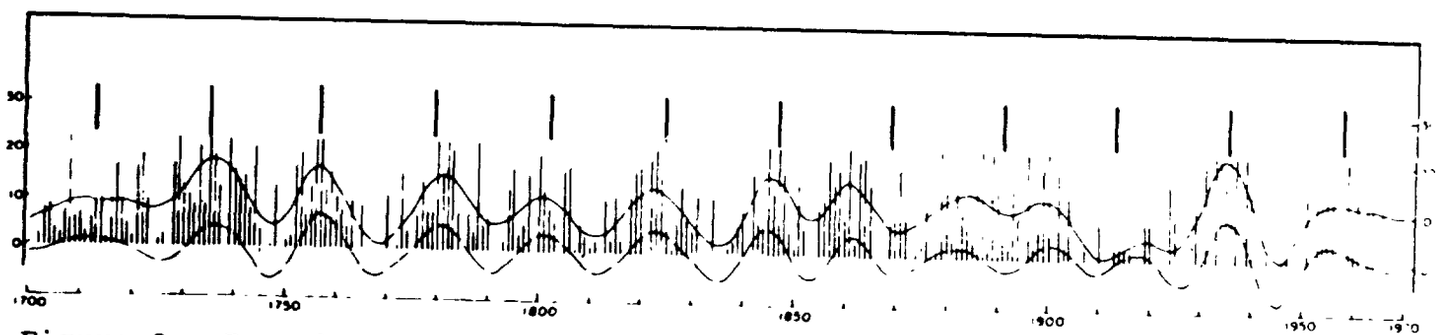


Figure 3. Drought area indices from tree ring records for the Western U.S. High Plains from Mitchell, Stockton & Meko (16). Wavy lines are their results from filtering the raw data with a band pass filter (lower) and a low pass filter (upper). The series of time marks above mark the regular phase corrected Hale magnetic sunspot cycle at 22.279 yr. intervals. Note the rapid phase shift from 1860-1900.

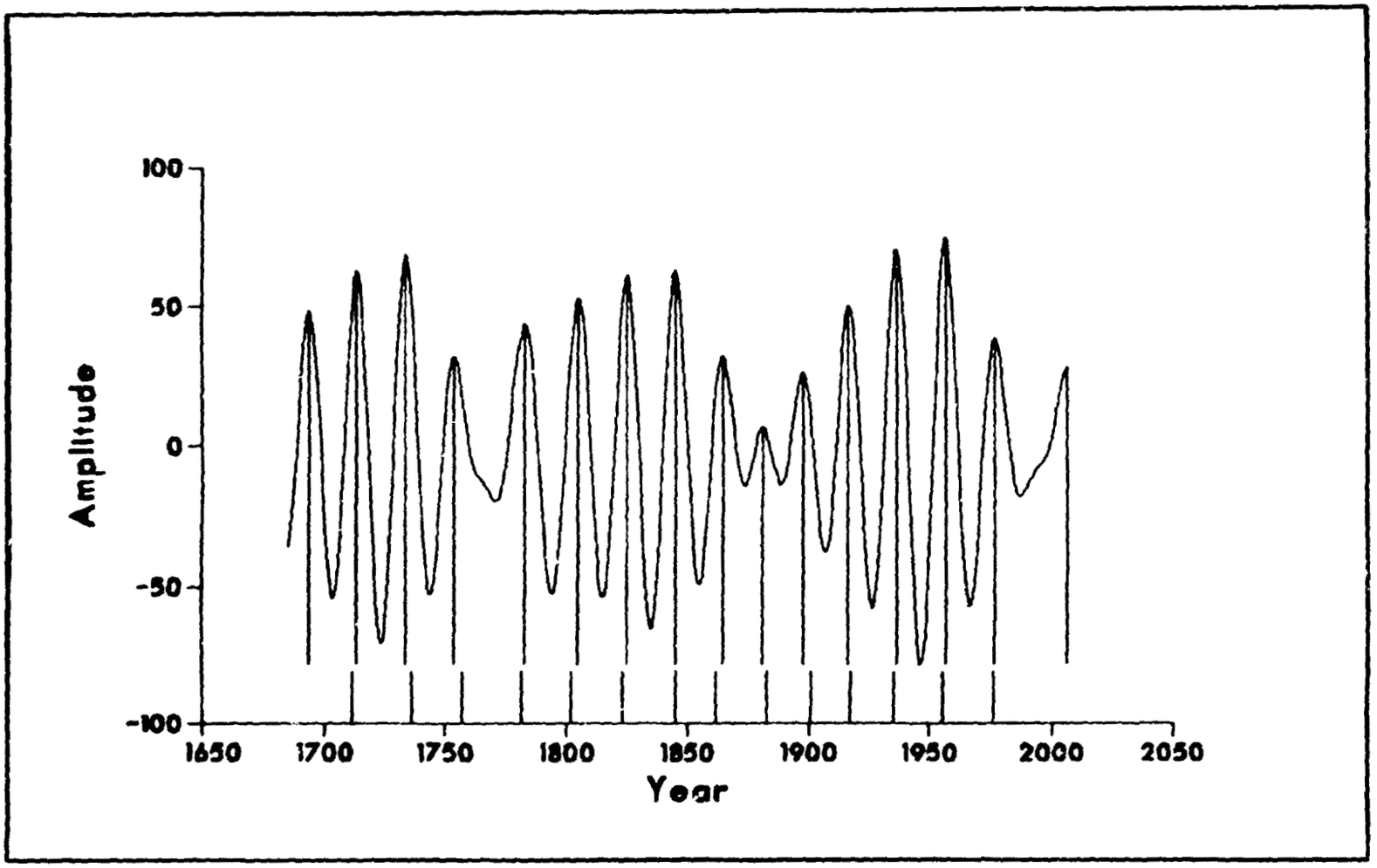


Figure 4. Beat wave between a constant 18.61 year period representing the lunar nodal tide cycle and the 22.279 year corrected Hale magnetic sunspot cycle with amplitude adjusted to fit the recorded sunspot activity. The long vertical lines mark the peaks of the beat wave, the short marks below them mark the times of maximum drought from Mitchell, Stockton and Meko.

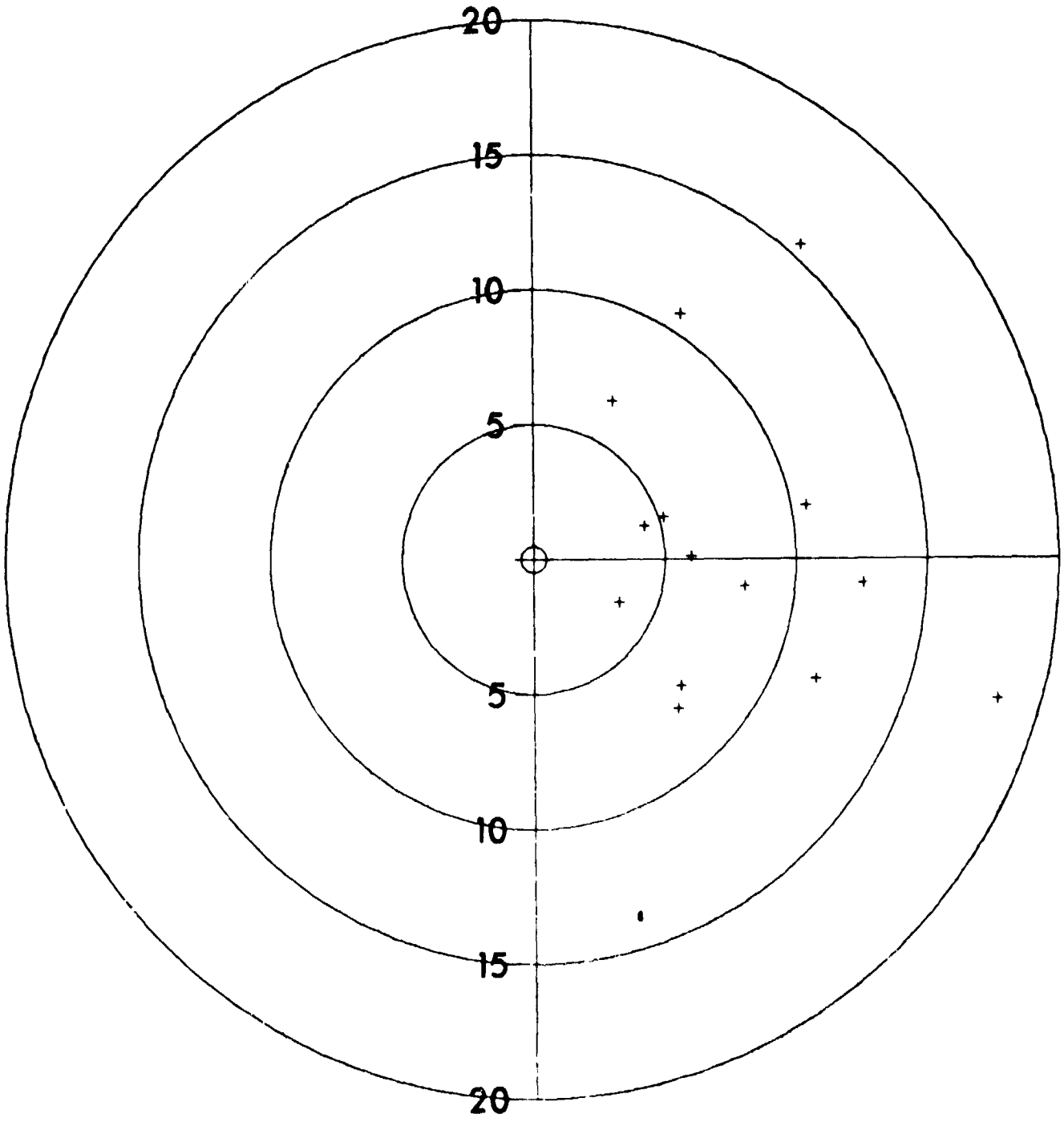


Figure 5b. Harmonic dial showing the phase relation between the solar cycle/lunar tidal cycle beat seen in Figure 4 whose peaks time the dial's rotation and the average of the drought times in Figure 5a. Note that now there are no large phase errors.

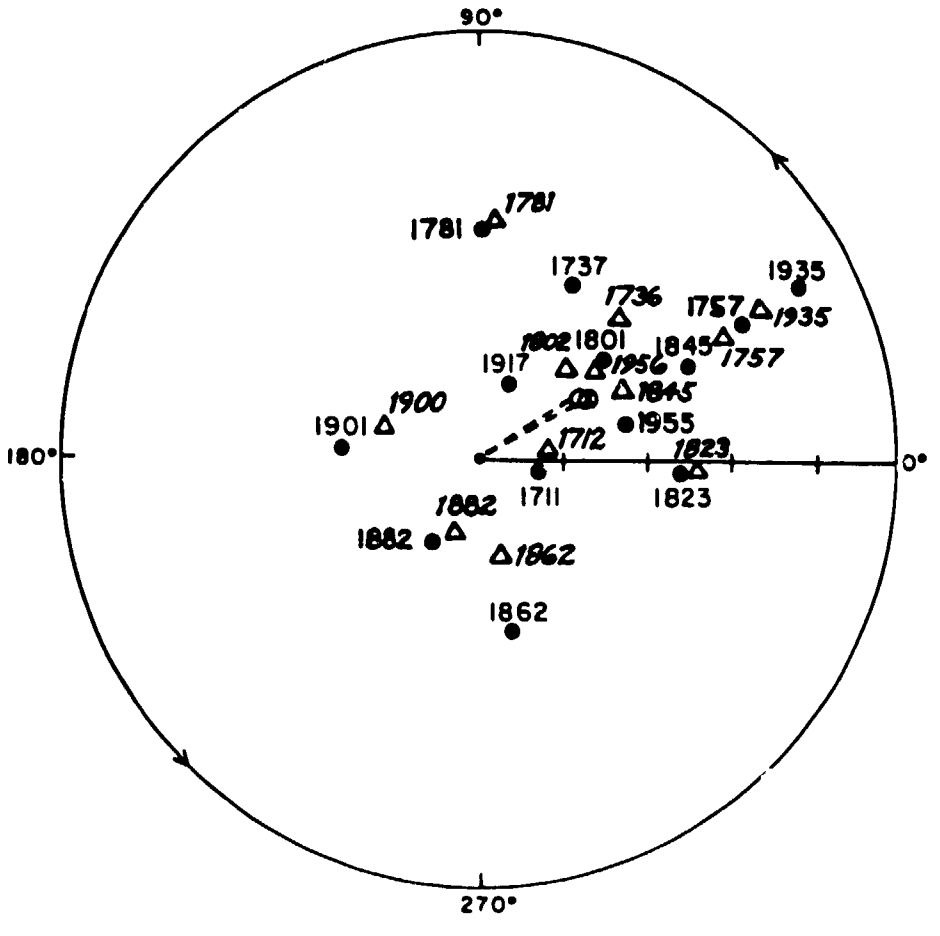


Figure 5a. Harmonic dial from Mitchell, Stockton and Meko (16) showing the phase relation between the uncorrected Hale magnetic sunspot period which times the dial's rotation and the drought area index series filtered by a band pass filter at 20.6 years (dots) and 24.3 years (triangles). Note the large phas errors at 1862, 1882, and 1900-1.

PREDOMINANT PERIODS IN THE TIME SERIES OF DROUGHT AREA INDEX FOR THE WESTERN HIGH PLAINS AD 1700 TO 1962

P. R. Bell
Institute for Energy Analysis
Oak Ridge, Tennessee

ABSTRACT

The detection of the combined presence of the Hale magnetic sunspot cycle (22.28 years) and the lunar nodal tidal cycle (18.61 years) in both the Eastern North American January air temperatures and the Western U.S. High Plains drought series has led to an extended analysis of the Drought Area Index time series. This analysis indicated that the mean dominant period of the drought series should be 20.0 to 20.5 years and that the principal period should be resolvable into two components of about 22.28 and 18.61 years. This note details the successful accomplishment of this task.

INTRODUCTION

The historical record of droughts in the United States Western Plains has been extended by the use of tree ring records by Mitchell, Stockton, and Meko (1), hereafter designated MSM. The tree ring data for 40 areas were evaluated from AD 1700 to the end of the tree ring library in 1963. This time series was calibrated to correspond to the instrumental Palmer Drought Severity Index (PDSI) in the time interval of overlap, between 1931 and 1963 (1). MSM studied this calibrated time series, called the Drought Area Index (DAI), for periodicities using discrete Fourier transform analysis. They interpreted the spectra as indicating a principal period of 22 years. They also performed filter analysis of the series corresponding to Palmer Drought Severity -1 (mildly dry) or worse (up to -4 [extremely dry]) with broad bandpass filters with center periods of 20.6 and 24.3 years and with a low pass filter. All the filter results were quite similar, showing a series of rather regular undulations. The timing of the peaks is an indication of periods of drought and the amplitude of the waves is an indication of drought intensity. The timing of the waves shows a good correlation with the Hale magnetic sunspot cycle and the amplitudes show good correlation with the envelope of the Zurich sunspot number (an indicator of general solar activity).

Recently Bell (2) has re-examined these results and the spectral and filter studies of the January mean surface air temperatures in Eastern North America (Mock and Hibler [3]). Bell demonstrated the presence of a beat note effect in both time series. This beat effect indicates the presence of two oscillations of comparable frequency and amplitude. The two periods found in the DAI series--by considering the gross period of the undulations and the number of waves between the two beat oppositions--were about 22.21 and 18.51 years. Bell suggested these periods were the 22.279 year Hale magnetic sunspot period determined by Dicke (4) and the 18.61 year lunar nodal tide cycle. The gross period, determined by counting the waves and the time span they covered (2), of the oscillations in the filtered DAI series of MSM was about 20.36 years.

It seemed worthwhile to reexamine the DAI time series using the higher resolution Maximum Entropy Spectrum Analysis (MESA) method to see if the two component periods could be resolved. Also, the DAI series constructed by MSM forms an amplitude scale that is quite non-linear since each of the 40 areas of the High Plains is counted as having either drought condition -1 or worse or else no drought for each year. In this way if the drought is more severe in an area it does not affect the total index for that year beyond 1 unit. We thought that the non-linearity of response might have biased the contribution of the component periods and distorted the spectral results.

DATA ANALYSIS

Through the kind cooperation of Dr. J. Murray Mitchell Jr., the Drought Area Index time series was made available for study. An effort was made to reduce the nonlinearity of the DAI series by forming a linear combination of the DAI indices -1 to -4. The value at the i-th year of the combined series is:

$$DAI(i) = A_{-1}*DAI\text{-}1(i) + A_{-2}*DAI\text{-}2(i) +$$

$$A_{-3}*DAI\text{-}3(i) + A_{-4}*DAI\text{-}4(i)$$

where the A's are weighting factors.

We do not know what values of the A's would produce a linear "drought" function, due as much to a lack of knowledge of a linear scale of "drought" as to the relation of drought scale to the Palmer Drought Severity Index from which the DAIs were calibrated. In our ignorance we

let all of the A's equal 1. Although this choice is quite arbitrary, it gives a better approximation to a linear scale since the value for an area may now increase as drought becomes more severe in that area.

A new time series was formed from the MSM series in this manner. To remove some of the high frequency noise, the series was then smoothed slightly by two passes of 1-2-1 smoothing, that is, for each pass:

$$DAISS(i) = [DAIS(i-1) + 2*DAIS(i) + DAIS\ (i+1)]/4$$

The mean value of the series was then removed from each DAISS(i) value. The resulting series is shown in Figure 1. Similar smoothing and average removal were applied to the original DAI-1 series of MSM.

Both of these time series were subjected to MESA. The results from both were quite similar. When short predictive filter lengths were used (L=24 to L=50) the principal peak was single-valued. At a filter length of 40 (Figure 2), the peak had a period of 20.04 years for the original DAI series and 20.23 years for the "more linear" DAISS series. At a filter length of 60 elements the principal peak was resolved into two components for both series. Figure 3 shows the spectrum for the "more linear" series at a filter length of 60, while Figure 4 shows an expanded section around the principal peaks. The periods of 22.07 and 18.60 years were found to be in good agreement with the expected values of 22.28 and 18.61 years (2). The original DAI series at a filter length of 63 gave component values of 22.03 and 18.23 years.

The filter lengths needed for resolution lie well below both the appropriate length for a short or noisy series at N/2(=131) recommended by Ulrych and Bishop (5) and the length for longer series at 2N/ln(2N)(=84) recommended by Berryman (6).

CONCLUSION

In conclusion, the high resolution MESA spectra provide evidence that the DAI time series is dominated by two principal oscillations that seem to correspond to the Hale magnetic sunspot cycle and the lunar nodal tidal cycle, in agreement with the findings of Bell (2). However, there is only a small change in the periods as the linearity of the DAI series amplitude scale is improved. The spectra (Figures 3 and 4) show that the lunar tidal cycle is considerably weaker than the solar cycle, on the average, but as a result of the considerable variability of the sunspot envelope, the amplitude of the lunar cycle was slightly greater at the beat opposition in 1880.

REFERENCES

1. Mitchell, J. Murray, Jr., C. W. Stockton, and D. M. Meko, in Solar-Terrestrial Influences on Weather and Climate, B. M. McCormac and T. A. Seliga, Eds. (Reidel Pub. Co., Holland, 1979), pp. 125-143.

2. Bell, P. R., "The Combined Solar and Tidal Influence in Climate," Workshop on Solar Constant Variations, NASA CP-2191, 1981.

3. Mock, S. J., and W. D. Hibler, III, "The 20 Year Oscillation in Eastern North American Temperature Records," Nature, 261: 484-486, June, 1976.

4. Dicke, R. H., "Solar Luminosity and the Sunspot Cycle," Nature, 280: 24-27, 1979.

5. Ulrych, T. J. and T. N. Bishop, "Maximum Entropy Spectrum Analysis and Autoregressive Decomposition," Rev. Geophys. and Space Physics, 13, 183-200, 1975.

6. Berryman, J. G., "Choice of Operator Length for Maximum Entropy Spectrum Analysis," Geophysics, 43(7): 1384-1391, Dec., 1978.

I gratefully acknowledge the helpful discussions with Dr. J. Murray Mitchell regarding the analysis of the data and his permission to analyze the tree ring drought data series.

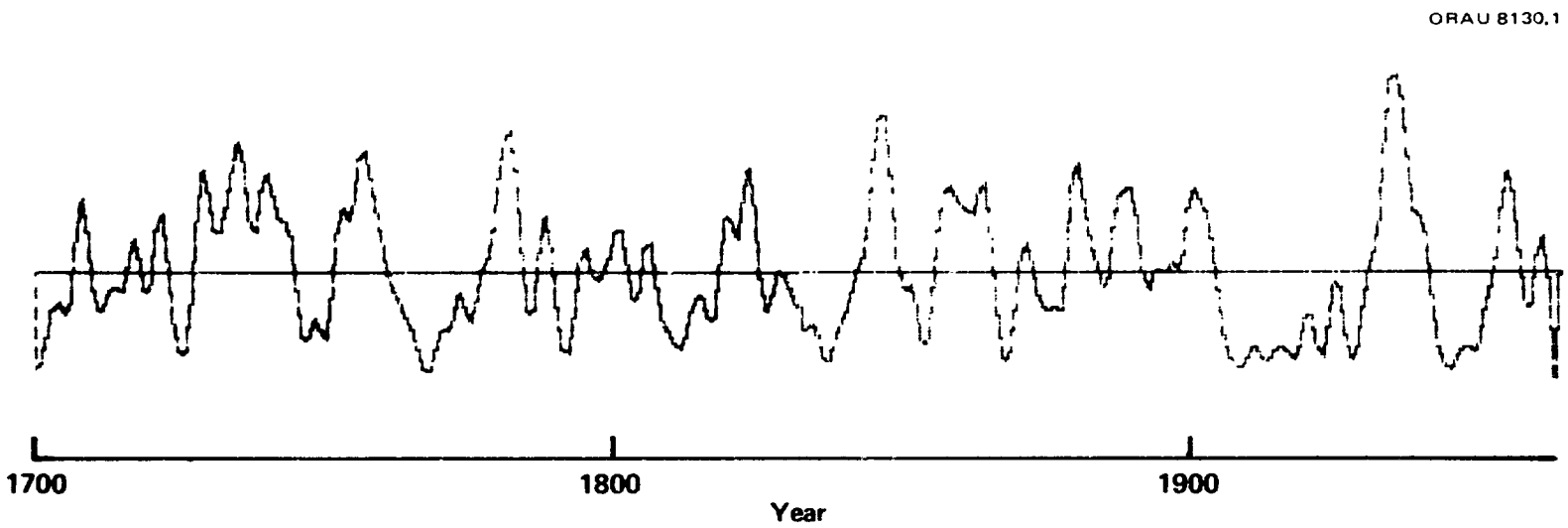


Figure 1. A more linearized Drought Area Index series for those areas out of 40 that had drought from mildly dry to worse from 1700 to 1962. The series has been smoothed by two passes of a 1-2-1 smoother.

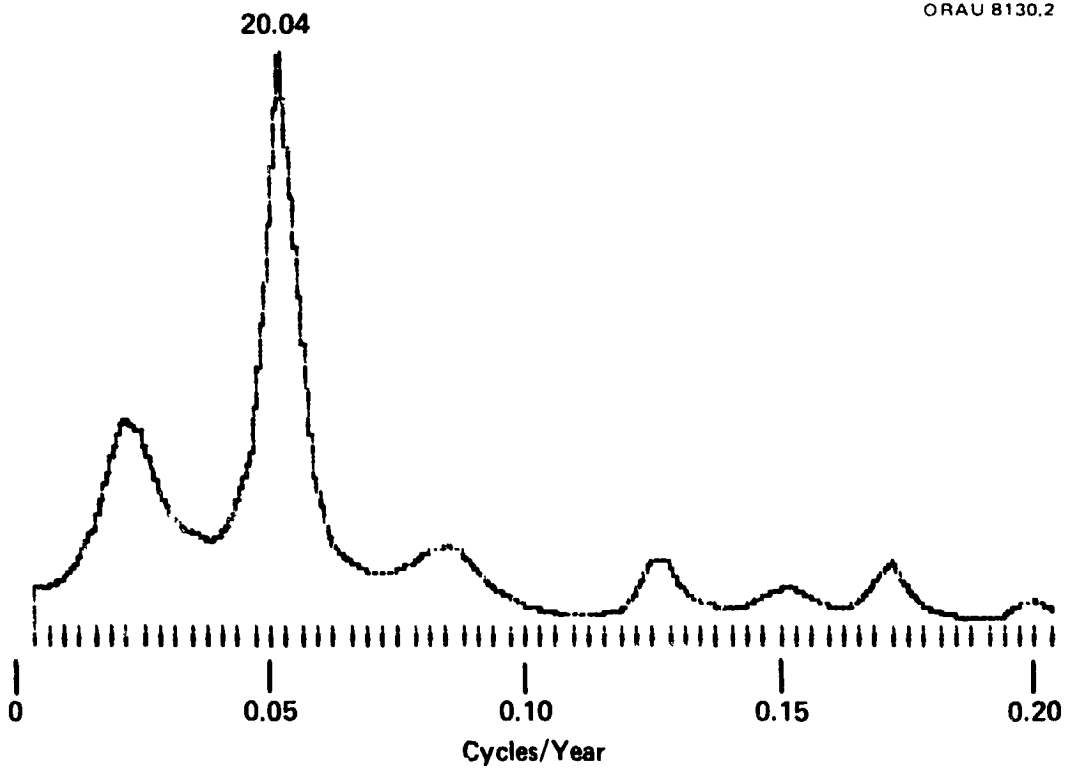


Figure 2. Maximum Entropy spectrum of the original DAI series at a filter length of 40 elements. The principal peak has a period of 20.04 years.

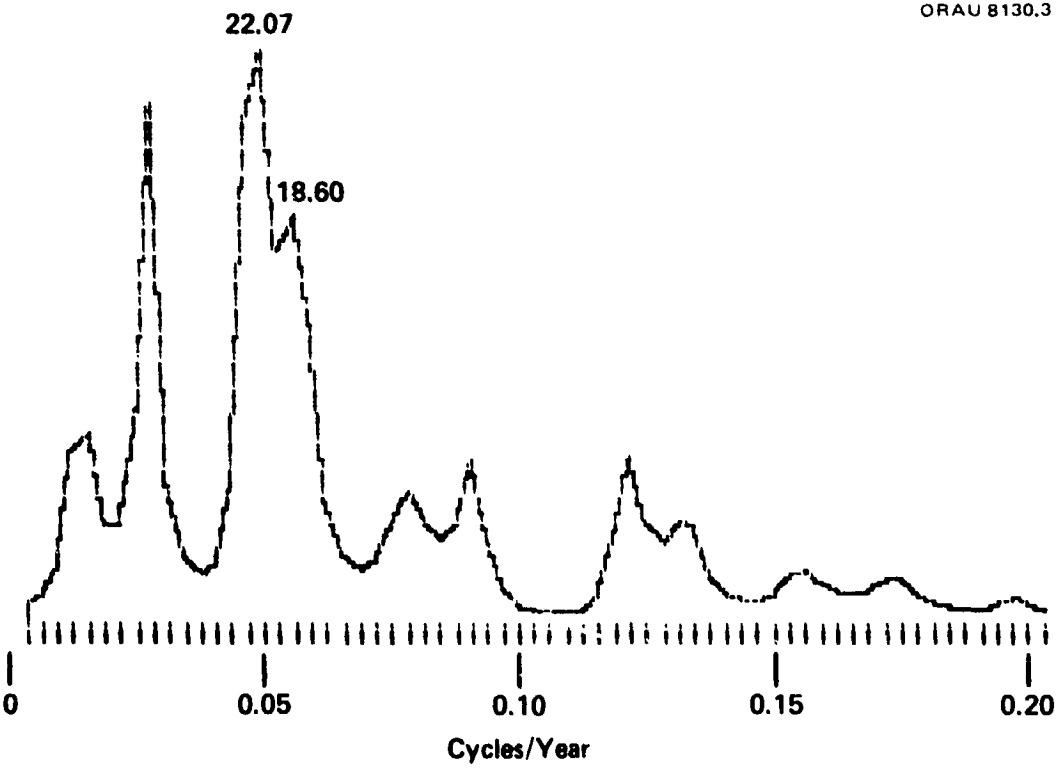


Figure 3. The MESA spectrum for the "more linear" DAI series at a filter length of 60 elements. The principal peak has been resolved into its two components of 22.07 and 18.60 years in good agreement with the 22.28 year Hale magnetic sunspot cycle and the 18.61 year lunar nodal tide cycle.

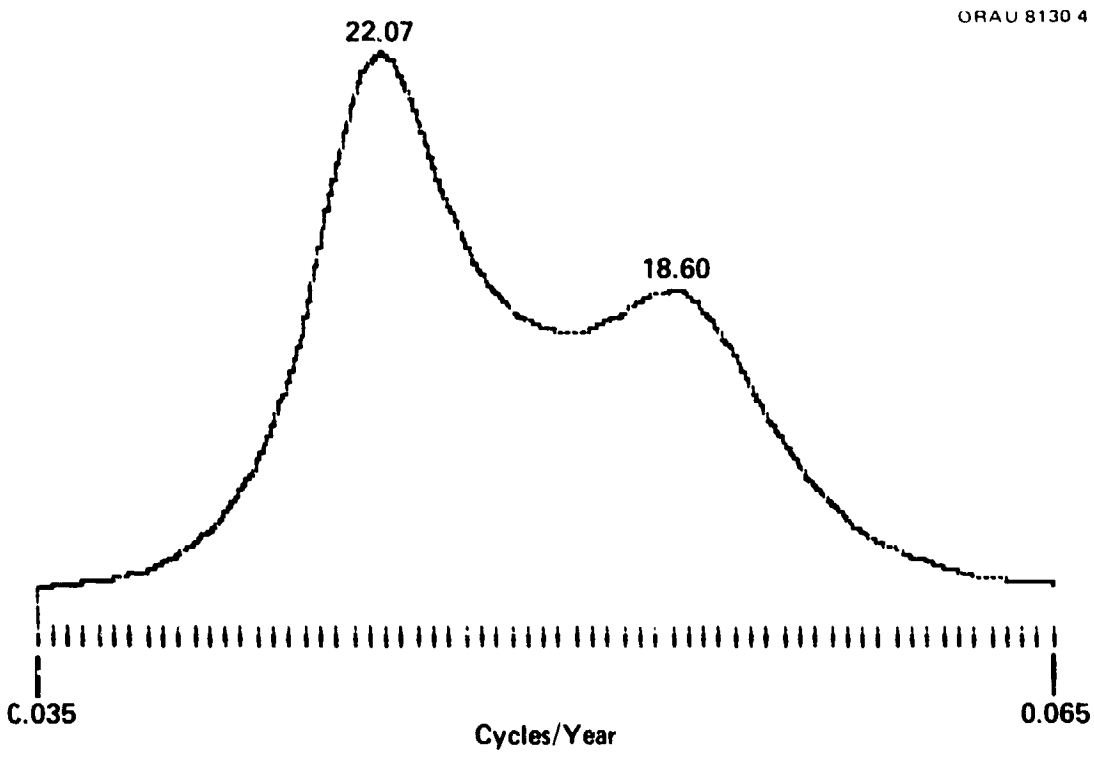


Figure 4. A section of the MESA spectrum of Figure 3 expanded

VARIATION OF THE SOLAR HE I 10830 Å LINE: 1977-1980

J. W. Harvey
Kitt Peak National Observatory*

ABSTRACT

Daily measurements of the equivalent width of the 10830 Å He I line integrated over the visible disk show:

1. An increase from about 32 to about 74 mÅ in the monthly mean values from the minimum to the maximum of the current solar cycle.

2. The monthly mean values are more smoothly varying than most other indices of solar activity.

3. Rotation modulates the daily values in a highly variable manner with amplitudes as large as ± 20%.

4. The apparent synodic rotation period is 29 days rather than the expected 27 days associated with active regions.

5. Despite great differences in the appearance of the sun in 3933 Å Ca I and 10830 Å He I, the central intensity of the former correlates with the equivalent width of the latter with a value r = 0.97.

INTRODUCTION

The 10830 Å line of He I is controlled in part by coronal radiation shining on the chromospheric layers where the line is formed (ref. 1). It is thus a convenient probe, readily observed from the ground with good spatial resolution (ref. 2), of coronal features such as coronal holes, bright points, filament channels, the bases of coronal streamers and other solar features which involve high temperatures (ref. 3). The line can be observed in a wide range of stars (ref. 4) and it is of considerable interest to ask how strong the line would appear in the sun, observed as a star, and how solar activity would be manifest in such observations. To address this problem, daily full disk spectroheliograms taken at Kitt Peak using the 10830 Å line have been

* Operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

spatially integrated to yield measurements of the equivalent width of the line in the sun seen as a star. The observations span a range from 1974 to the present but only 1977 through 1980 have been completely processed.

RESULTS

SOLAR CYCLE VARIATION

Figure 1 shows the variation of the 10830 Å line equivalent width averaged over one month intervals from 1976 to 1980. The low value of about 32 mÅ corresponds in time with the minimum of solar activity while the high value of 74 mÅ corresponds to the peak of solar activity in November 1979. This plot has been compared with monthly mean sunspot numbers, calcium plage index, 2800 MHz flux values, soft x-ray flux values, etc. The best correlation is with 2800 MHz flux but the 10830 Å values vary significantly more smoothly than this and other indices. Presumably this behavior is due to the relatively large contribution to the 10830 Å signal from old, long-lived active regions.

ROTATIONAL MODULATION

Plots of daily 10830 Å values covering a long time period show a clear signature of solar rotation. The typical behavior is a period of obvious modulation for 3-5 rotations followed by a similar period of little discernible rotational modulation. The amplitude of the modulation has been observed to be as large as ± 20% of the mean values. Figure 2 shows an example of strong rotational modulation. The cause of this modulation can be found by examination of the spatially resolved daily images. The modulation is due to active regions clustered in a limited range in longitude for a period of a few months. Once in a while the clustering of active regions produces modulation at half the rotation period as, for example, in figure 3.

A surprise was the rotation period derived from the plots. It is found to be 29 days synodic rather than the 27-day synodic period expected for active regions. The cause of this discrepancy appears to be successive development of active regions at the eastern edge of an active longitude range together with rapid decay of active regions at the western edge of the range. The individual active regions within a limited range of active longitudes do rotate with the expected period of slightly more than 27 days. Whether or not this behavior in active region development indicates some large scale phenomenon requires further study.

COMPARISON WITH CALCIUM MEASUREMENTS

Images of the sun made with the 10830 Å He I line do not look like solar

images made using other spectral lines. In particular, the helium images show filaments as the strongest features, a rather diffuse appearance to active regions, and a manifestation of coronal features such as coronal holes and bright points. Therefore if one compares daily integrated flux measurements of the sun made with 10830 Å and say 3933 Å Ca II, one would expect considerable differences. This comparison is shown in figure 4 based on daily measurements of the central intensity of the calcium K line by White and Livingston (ref. 5). Contrary to expectation, the agreement is very good with a linear correlation coefficient r = 0.97. Evidently active regions dominate the 10830 Å signal (as they do for the K line) so that other 10830 Å solar features which do not appear in the K line are feeble by comparison.

CONCLUSIONS

The 10830 Å He I line provides a satisfactory means for monitoring solar cycle variations of the chromosphere. In spite of its sensitivity to coronal radiation, the variation of the 10830 Å line strength with time closely follows other indices of solar activity, especially the 2800 MHz and K line integrated fluxes. The unique value of the 10830 Å line appears to be in measuring stellar cycles in other stars in which the K line is difficult to observe.

Experience with the sun suggests that use of chromospheric line fluctuations as indicators of stellar rotational periods is not straightforward. The rotation period we derive differs from the true rotation rate by 7%. This curious discrepancy may be due to a large-scale ordering of the occurrence of solar activity, and may therefore prove to be a powerful diagnostic of such a large scale phenomenon, but more work is needed.

It is surprising and somewhat disappointing that the 10830 Å line measurements agree so well with K line measurements. This means that high temperature phenomena in the solar atmosphere cannot be monitored easily with the 10830 Å line. Presumably the same difficulty would arise in monitoring other stars. The result is that it may prove to be very difficult to discern important features such as coronal holes in other stars by ground-based measurements.

REFERENCES

1. Goldberg, L.: The Temperature of the Solar Chromosphere. Ap. J., 89, 1939, pp. 673-678.

2. Livingston, W. C.; Harvey, J.; Pierce, A. K.; Schrage, D.; Gillespie, B.; Simmons, J.; and Slaughter, C.: Kitt Peak 60-cm Vacuum Telescope. Appl. Opt., 15, 1976, pp. 33-39.

3. Harvey, J. W.; and Sheeley, N. R., Jr.: A Comparison of He II 304 Å and He I 10830 Å Spectroheliograms. Sol. Phys., 54, 1977, pp. 343-351.

4. Zirin, H.: Further Observations of the λ10830 Helium Line in Stars and their Significance as a Measure of Stellar Activity. Ap. J., 208, 1976, pp. 414-425.

5. White, O. R.; and Livingston, W.: Solar Luminosity Variation. II. Behavior of Calcium H and K at Solar Minimum and the Onset of Cycle 21. Ap. J., 226, 1978, pp. 679-686.

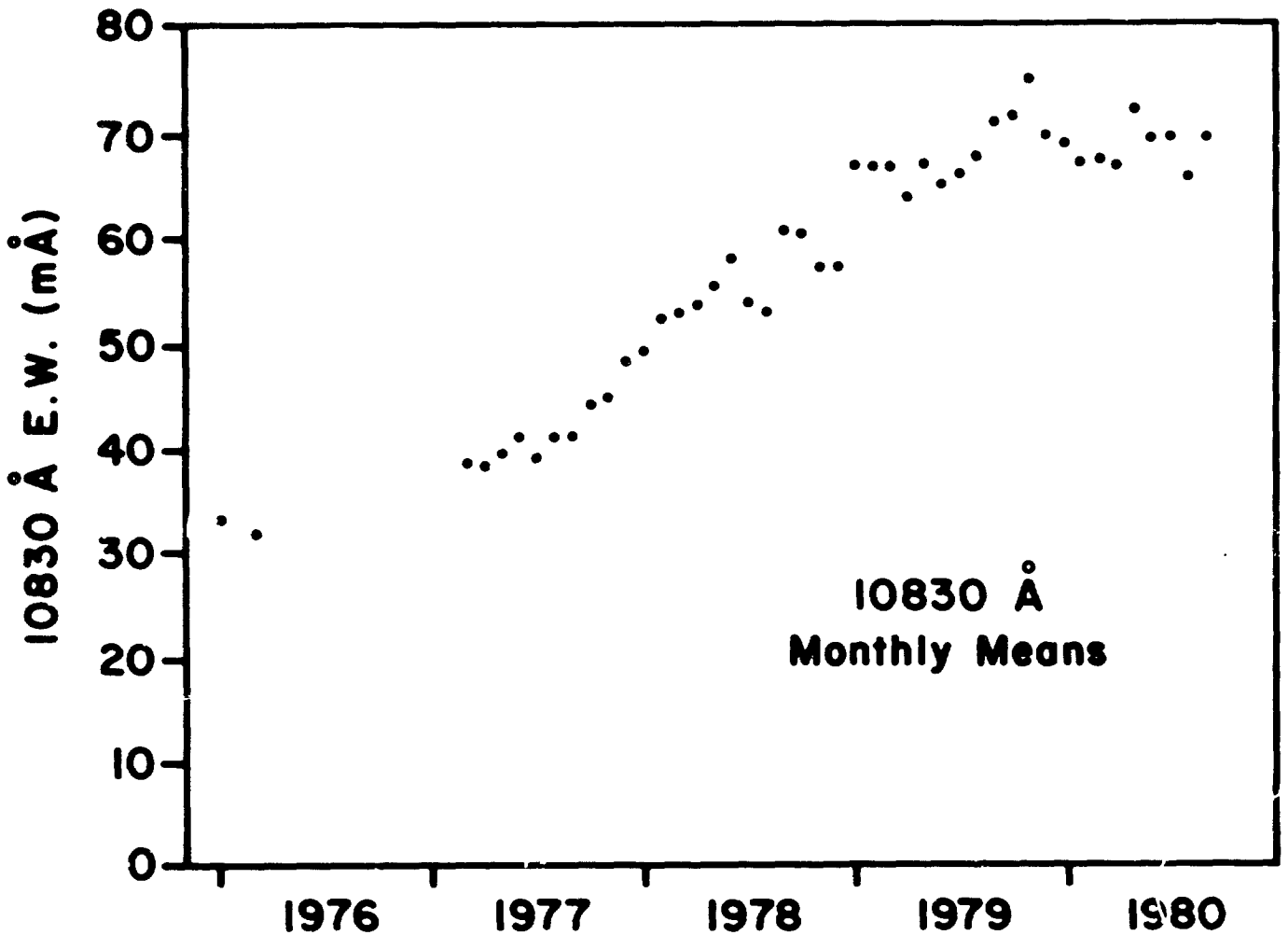


Figure 1. Monthly mean values of the equivalent width in mÅ of the 10830 Å line integrated over the solar disk. The values rise from solar activity minimum in 1976 to the maximum in late 1979.

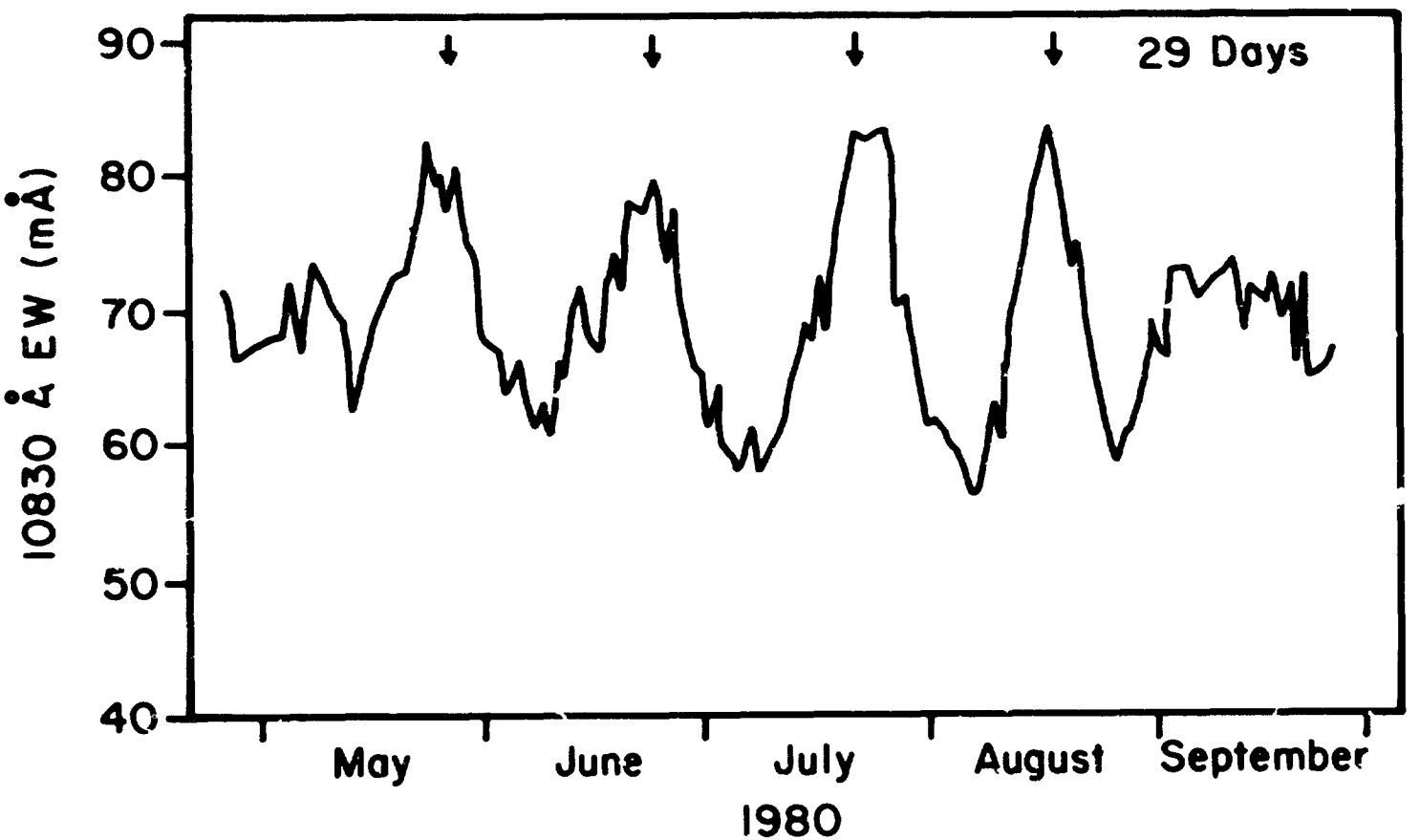


Figure 2. Daily values of the equivalent width in mÅ of the 10830 Å line integrated over the solar disk for a period of strong rotational modulation of the signal. Arrows indicate a synodic recurrence period of 29 days rather than the expected 27 days.

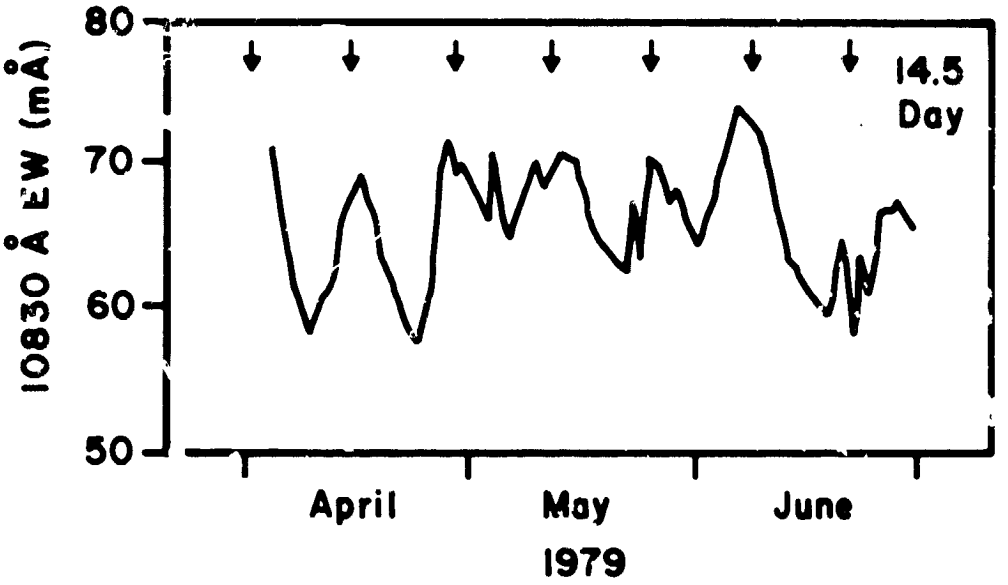


Figure 3. Daily values of the equivalent width in mÅ of the 10830 Å line integrated over the solar disk for a period when the rotational modulation was dominated by a 14.5 day period rather than the 27 day period normally present.

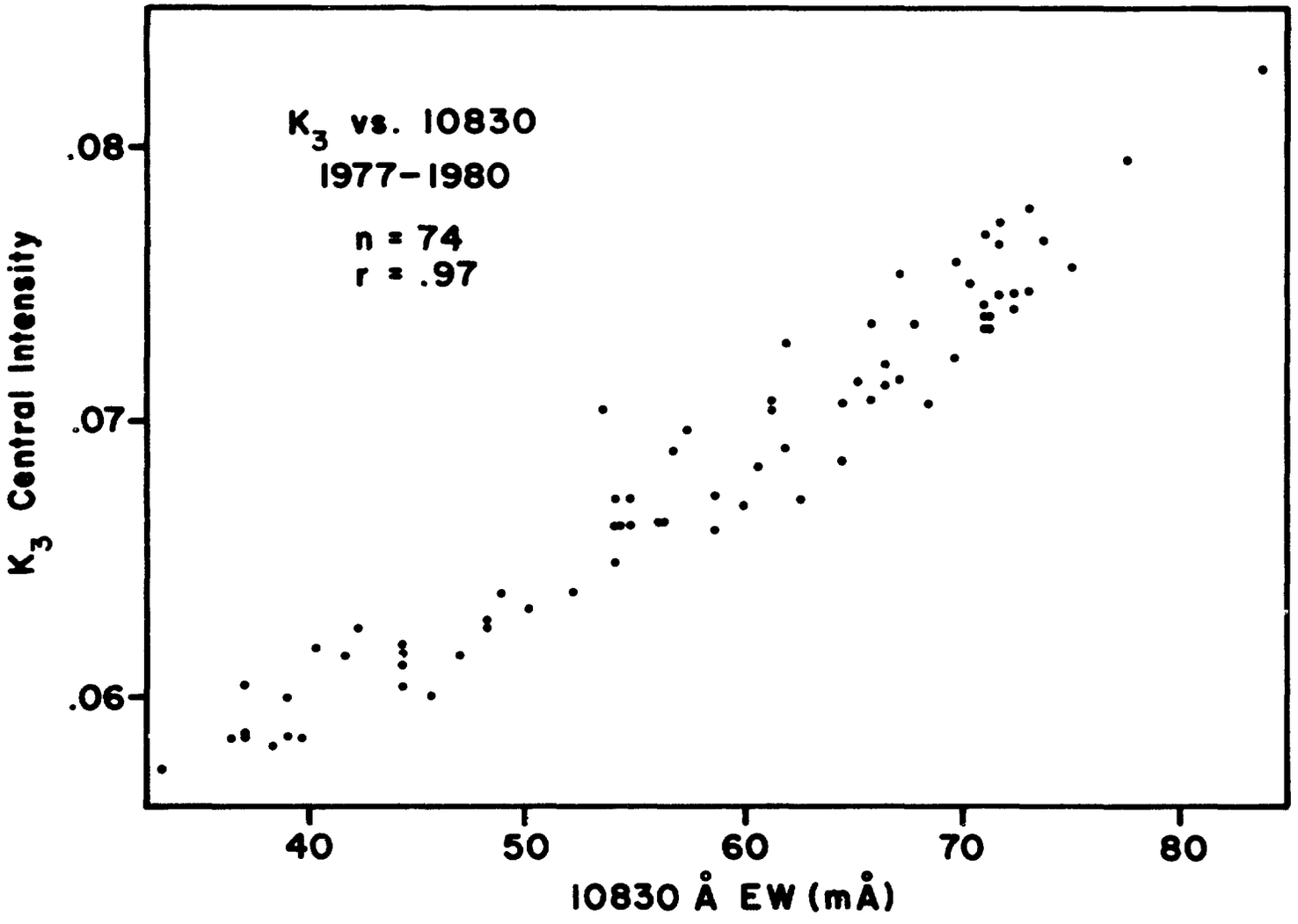


Figure 4. Scatter plot of the central intensity of the calcium K line and the equivalent width of the 10830 Å line. Each point represents observations made within a few hours of each other on 74 different days from 1977 through 1980.

A PROSPECTUS FOR A THEORY OF VARIABLE VARIABILITY*

S. Childress
Courant Institute of Mathematical Sciences
New York University

E. A. Spiegel
Astronomy Department, Columbia University

ABSTRACT

We propose that the kind of stellar variability exhibited by the sun in its magnetic activity cycle should be considered as a prototype of a class of stellar variability. The signature includes long 'periods' (compared to that of the radial fundamental mode), erratic behavior, and intermittency. As other phenomena in the same variability class we nominate the luminosity fluctuations of ZZ Ceti stars and the solar 160^m oscillation. We discuss the possibility that analogous physical mechanisms are at work in all these cases, namely instabilities driven in a thin layer. These instabilities should be favorable to grave modes (in angle) and should arise in conditions that may allow more than one kind of instability to occur at once. The interaction of these competing instabilities produces complicated temporal variations. Given suitable idealizations, we show how to begin to compute solutions of small, but finite, amplitude and we discuss the prospects for further developments.

THE PROPOSAL

VARIABLE VARIABILITY

An aim of this paper is to argue that the kind of variability that the sun displays in its magnetic activity cycle is the prototype for a category of stellar variability that should be isolated and studied as a generic phenomenon. We are not referring here to the group of solar type stars that show magnetic activity, though they and the sun do make up a class of variable star in the usual sense of the term. We are speaking of a kind of variability and thus of a broader category, if a more abstract one, than that of a class of star. The kind of variability that we have in mind includes intermittency, such as the sun exhibited

* The financial support of this work by the National Science Foundation under Grants NSF PHY-77-27086 at Columbia and NSF MCS-79-02766 is gratefully acknowledged.

when it went through the Maunder minimum [1], and irregularity such as the cycle displays when it is detectable. We presume that all solar type variables show this variable variability, as we may call it. We suggest that the time dependence of ZZ Ceti stars [2] is of the same kind, details aside. The stars in this latter class are variable white dwarfs with marked temporal intermittency and, if they are in the same variability category as the sun, this could be useful for solar studies since ZZ Ceti periods are of the order of ten minutes. We believe too that the sun shows variable variability in several ways and in particular through the 160^m oscillation, which is sometimes quite hard to detect and is always very noisy [3]. There are other examples that come readily to mind, but we mention only these since they are the only ones that we have done serious calculations for in the way that we shall describe below.

The point of isolating a kind of variability is that it may help to identify the physical mechanism that produces the variations. In the case of the simplest kind of stellar variability, that exemplified by the regularity of Cepheids, we normally try to find a mechanism analagous to that of the Cepheids, in which case we might expect that a static star is overstable or vibrationally unstable to small perturbations. Then periodic solutions will bifurcate from the static solutions in a fashion called a Hopf bifurcation [4] nowadays. Overstability is more complicated than ordinary or direct or dynamical instability, in which perturbations to a static configuration grow monotonically and steady solutions bifurcate from the static ones.

MODELS OF APERIODICITY

The contrast between direct instability and overstability is made vivid by systems that can manifest either depending on the value of some system parameter such as angular velocity. We suggest that such a system is implicated in the solar cycle. Qualitative evidence for this remark is provided by a model constructed to explain the solar 5^m oscillation as a convective overstability [5]. The model could be either overstable or directly unstable according to the values of certain parameters. As in most such systems there are two important parameters, one controlling the amount of each instability. The system is said to have codimension two [6]. Such a system can be rewired, so to say, so that one parameter controls the relative amounts of both kinds of instability and another controls the total amount of instability. When the system is set to hover between the two instabilities and the amount of instability is turned up, the resulting oscillations become aperiodic. This led to the conjecture that unstable systems hovering between the two kinds of instability would generally behave aperiodically. This suggestion was tested on a one-zone model for radial pulsation [7]. The idea was that for a mean $\Gamma < 4/3$ one gets direct (or dynamical) instability while for $\Gamma > 4/3$ one may have overstability. For $\Gamma \approx 4/3$ one might expect

erratic behavior, and that was what was found. This is the kind of behavior we are postulating to explain the time dependence of the solar cycle and of other variable variables.

We have to ask at once whether it is reasonable to expect to find many systems arranged not only to be unstable but also to be unstable in several ways. Our answer is yes but our reasons are complicated. Let us simply say here that one reason lies in the circumstance that many systems do have long periods compared to their natural radial periods. To see what this may mean, consider the case of convective overstability. How may we engender slow variations in the elementary theory of convection?

CELL SIZES AND TIME SCALES

When you heat a fluid from below to induce convection you generally are doing something thermally complicated, but approximately you are usually fixing the temperatures on the top and bottom boundaries of the fluid. When the imposed temperature difference is large enough, sustained motions may begin [8]. These motions are organized into cells which tesselate the layer on a horizontal scale comparable to the layer thickness. However, if you instead specify the heat flux on the boundaries, the cells are as large in the horizontal as the geometry will allow [9]. Such big cells are easier to excite and are slower to react than the more popularly sized cells.

In a case where overstability is also possible, the role of the boundary conditions is very important too. For, not only are the growth rates small, as in normal convection, but the frequencies of the overstable modes are also small. Hence boundary conditions that favor large horizontal scales in rather thin layers will tend to put such systems willy-nilly into the states we want them. Whether those boundary conditions are realistic in a given configuration cannot be stated in advance, but at least they are frequently not implausible. Moreover, it is not always easy to say ahead of time precisely which conditions will induce large horizontal scales and relatively low frequencies. But that they may arise naturally is attested by studies of instabilities relating to ZZ Ceti stars [10,11]. The periods turn out to be long compared to the fundamental period of the star which is reckoned in seconds. This is just the kind of situation that we need for the analysis we shall describe.

GETTING STARTED

Even if it is true that mild instabilities in thin shells lead to the variable variability we have described, it might not always be obvious which thin shell is involved. However once the idea is there, we have the motivation to look for the right kind of instability. For the solar cycle a clue is provided by the familiar problem [12] of building strong magnetic fields in many

solar dynamo models. The general idea is that differential rotation produces a toroidal field whose strength builds till it goes unstable and buckles to protrude from the sun and create spots. The difficulty in this picture is that turbulent convection and magnetic buoyancy [12] will quickly destroy any ordered field so it is not clear how a significant toroidal field can persist for any time in the solar convective zone. A way out of this difficulty is to form the ordered field just below the convective zone [13]. Large scale convection produces flux expulsion [12] and topological pumping [14]. However, since the material just below the convection zone is a good conductor, it does not readily allow the field to penetrate it. The field is nevertheless swept into an intermediate region by penetrative convection, which may be mild enough to leave it ordered. As the field builds up in the bottom of the penetrative zone, the penetrative motion will be impeded and the convection zone recedes leaving behind a layer of ordered field. Ultimately, the layer becomes thick enough for magnetic buoyancy to overwhelm the local stable temperature gradient. There follows a new round of solar activity. Of course other observable manifestations of such a process should exist and it is unclear as yet whether we are on firm grounds [15].

THE 160^m OSCILLATION

For another illustration of the procedure consider the 160^m oscillation. Here there is no accepted explanation and our view is that the nature of the oscillation leads us to look for a thin overstable layer. The low frequency points to gravity waves and perhaps to an important role for buoyancy. The popular objection to explaining the 160^m period as owing to gravity waves is that the spectrum should be dense and some claim that a broad band of frequencies should be excited. However, if we are dealing with a thin layer, we may study waves whose lengths are much greater than the layer thickness. It is then possible to strike a balance between nonlinearity and dispersion so that a solitary wave or a train of them is produced as in the theory of shallow water waves [16]. In this picture, which may be described by our procedures, the 160^m period should be the travel time of the solitary wave around the sun. This gives a clue to where the waves are excited.

In a stably stratified gas, linear gravity waves propagate with a maximum speed given roughly by [17]

$$c_g = \left(\frac{\gamma-1}{\gamma}\, gH\right)^{\frac{1}{2}} = \left(\frac{\gamma-1}{\gamma}\, R_* T\right)^{\frac{1}{2}},$$

where R_* is the gas constant. For present purposes we can write this as

$$c_g \sim 10^6 \sqrt{T_6}\ \text{cm/s},$$

where $T_4 = T/10^4 K$. If the layer in question has radius $7\times10^{10} r$ cm, we find a period

$$P = 10^4 \, r/\sqrt{T_4} \text{ min.}$$

So we want a layer such that

$$T_4 = 3600r.$$

This layer needs to be in the deep interior. We propose that the argument makes it worthwhile to inquire whether the 160^m oscillation can be excited in the solar core.

Fortunately, a suitable mechanism is already known: Dilke and Gough [18] found an overstability that should exist in the present solar core. For this instability to be in keeping with our prescription we must postulate a thin layer rich in ^3He at the edge of the nuclear burning region. Granted this, the kind of nonlinear analysis that we shall now illustrate goes through. It has been done for a mildly nonBoussinesq model and it needs more refinements. Nevertheless there are some interesting features of the solutions. For example, a mild thermal anomaly propagates about the core. This contributes to energy generation, and to get the right solar luminosity, we have to lower the central temperature a bit. How significant the change is depends on the amount of ^3He we put in; a careful comparison with the oscillation data will be needed to make a quantitative statement. In fact no numerical estimates of any kind are given in the next section, which is too physically bare to be anything but a demonstration model. As we shall explain below, it is too primitive mathematically. Yet we think its general design is good and if you are looking for an approach to these problems, you might want to consider this possibility.

THE PROCEDURE

FORMULATION OF A TRACTABLE PROBLEM

The General Model

To discuss the dynamics of a hypothetical magnetic layer located just under the solar convective zone, we might reasonably presume that this layer is subject to a given heat flux from below and that it passes this same flux on to the layers above it. Similarly, we might guess that the overlying convective zone pumps a horizontal magnetic field downward into the layer. In the static state the field continues to diffuse slowly downward, and there will be a slight time dependence. We may avoid this complication by assuming that the field is removed from the bottom boundary of

the layer at the rate that it is being fed in at the top. This is the idealization we propose for the sun's magnetic layer. For semiconvection we would likewise assume that the flux of helium at the top and bottom boundaries is prescribed.

We assume that the geometry is plane-parallel and let z be the upward coordinate, y be the horizontal coordinate corresponding to the eastward direction and x be the equivalent northward direction. The fluid is confined to the layer $-d/2 \leq z \leq d/2$.

Though the calculations that follow are detailed, they are stripped down to the barest essentials for this demonstration. No extras are included. Thus we assume here that all of the parameters characterizing the fluid, such as the gas constant, the acceleration of gravity, the specific heat at constant pressure, the permeability, and the several diffusivities, are constants. The main dependent variables of interest for the convective process are the velocity and the state variables such as temperature, magnetic field and molecular weight. To describe state variables we shall use a standardized notation which we illustrate for the case of temperature.

The Dependent Variables

Let the temperature, T, be decomposed into a static part and a convective part:

$$T(x,z,t) = \bar{T}(z) + \delta T(x,z,t)$$

where t is time. We allow no y-dependence and consider only a two-dimensional problem. Let

$$T_o = \bar{T}(0)$$

be used as a temperature scale and let

$$\xi_T = \frac{d}{\bar{T}_o}\frac{d\bar{T}}{dz}.$$

This measures the temperature contrast across the static layer. A scaled temperature disturbance is then defined by

$$\delta T = \left|\xi_T + \frac{g\,d}{C_p T_o}\right| T_o\,\theta,$$

where C_p is the ratio of specific heats and g is the acceleration of gravity. For the other state variables we proceed similarly. For example we might introduce a scaled magnetic perturbation in the two dimensional case where the field has only a y-component:

$$\delta B = \left|\xi_B - \xi_\rho\right| B_o\,\Sigma,$$

where ρ is density. An analogous scaled variable would be used for molecular weight.

We shall assume that the velocity is solenoidal, and this is not a bad approximation for the kind of gravity wave solution we are after. The velocity can therefore be described by a non-dimensional stream function:

$$\mathbf{u} = \frac{\kappa}{d}\left(-\frac{\partial \psi}{\partial z}\ ,\ 0,\ \frac{\partial \psi}{\partial x}\right),$$

where κ is the thermal diffusivity.

The Boussinesq Equations

Next we make an approximation which forces us to give up some qualitative features of the problem that are of astrophysical interest, but which permits us to bring out the basic nature of the phenomenon as simply as possible and to illustrate the kind of calculation that the subject entails. This is the Boussinesq approximation in which we omit density fluctuations except insofar as they directly produce driving by buoyancy forces. Boussinesq theory also neglects pressure fluctuations in the equation of state. The pressure fluctuation enters only as a gradient term in the equation of motion as needed to maintain the solenoidal condition on the velocity field. In hydromagnetic convection the analogous approximation is the neglect of fluctuations in the total pressure (gas plus magnetic). On introducing these approximations we are led to the following equations for two-dimensional Boussinesq magnetoconvection [19]:

$$(\partial_t - \sigma\nabla^2)\nabla^2\psi = -\sigma\zeta_\Theta R\partial_x\Theta + \sigma\tau\zeta_\Sigma S\partial_x\Sigma + \frac{\partial(\psi, \nabla^2\psi)}{\partial(x,z)}, \tag{1}$$

$$(\partial_t - \nabla^2)\Theta - \partial_x\psi = \frac{\partial(\psi,\Theta)}{\partial(x,z)}, \tag{2}$$

$$(\partial_t - \tau\nabla^2)\Sigma - \partial_x\psi = \frac{\partial(\psi,\Sigma)}{\partial(x,z)}, \tag{3}$$

where

$$R = \frac{gd^3}{\kappa\nu}\left|\xi_T + \frac{g\,d}{c_p T_o}\right|, \qquad S = \frac{gd^3}{\eta\nu}\frac{B_o^2}{\mu p_o}\left|\xi_B - \xi_\rho\right|,$$

$$\sigma = \nu/\kappa\ ,\ \tau = \eta/\kappa\ .$$

Here ζ_F is +1 if the vertical gradient of F is in the unstable sense and is -1 if it is in the stable sense (F may be Θ or Σ), p stands for gas pressure, ν for kinematic viscosity, μ for permeability, and η for magnetic diffusivity. The Rayleigh (R and S) and Prandtl (σ and τ) numbers appear because we have used natural units for length (d) and time (d^2/κ). We have followed one standard practice in assuming that the Rayleigh numbers are positive and

separating out the stability discriminants, ζ_F. We find this cumbersome and we prefer to adopt a particular choice for them:

$$\zeta_\Theta = 1, \ \zeta_\Sigma = -1,$$

which is the typical case for semiconvection [20]. We make this choice even though semiconvection plays a secondary role in destabilizing ZZ Ceti stars [10] because it is one of the few cases for which we know explicitly the right discriminants to select. For the magnetoconvective model of the solar cycle, a number of possible ζ combinations may arise in the important magnetic layer and, in any case, for that situation, there appear to be vital nonBoussinesq effects. (We suspect too that differential rotation is important and that adds another ζ.) So it would be misleading to try to describe a solar model at the present minimal level. But elementary as the present example is, it seems to contain the essential nonlinear dynamics of semiconvection, which may play a very significant role in certain variable stars and whose time dependence, we suspect, is a good example on which to base thinking about variable variability.

The Boundary Conditions

When we go from magnetoconvection to semiconvection, in two dimensions, we have to deal with precisely the same set of Boussinesq equations, but now Σ is the perturbation in molecular weight instead of the magnetic disturbance. These equations are also the right ones for thermohaline convection, which has a much greater following [21] and which has inspired our notation (Σ for salinity, S for saline Rayleigh number). All the computations we know of in this subject have been done with Θ and Σ vanishing on the upper and lower boundaries [24]. As advertised, we here fix fluxes on the upper and lower boundaries. Hence, the pertubation fluxes must vanish on these boundaries and we require that

$$\partial_z \Theta = 0, \ \partial_z \Sigma = 0 \quad @ \ z = \pm\tfrac{1}{2}. \qquad (4a,b)$$

We also need kinematic conditions on top and bottom and we assume that the boundaries are stress free but not deformable. Then

$$\psi = 0, \ \partial_z^2 \psi = 0 \quad @ \ z = \pm\tfrac{1}{2}. \qquad (5a,b)$$

An advantage of having the fixed-flux boundary conditions is that they generally favor large horizontal scales and once this is realized, we can use the methods generally associated with the theory of shallow water waves [16,22]. This permits us to include in our asymptotic studies several effects that have been difficult to treat other than by numerical methods. But the main physical interest of these large horizontal scales is that they also involve slow behavior such as we would like to postulate in a model for variable variability.

The Scaled Equations

To find approximate solutions we proceed by asymptotic methods, combining amplitude expansions with scaling of time and space coordinates. To ensure small amplitude, we assume that the layer is only mildly unstable. Let

$$R = R_0 + \varepsilon^2 R_2, \qquad S = S_0 + \varepsilon^2 S_2,$$

where $\varepsilon^2 \ll 1$ and R_0 and S_0 are values of R and S that re[illegible] fluid neutrally stable while R_2 and S_2 are arbitrary.

We anticipate that, as in ordinary convection theo[illegible] amplitude of the motion is $O(\varepsilon)$ and we therefore set

$$\psi = \varepsilon \Psi .$$

However it is best not to rescale the temperature in Boussinesq convection with fixed flux [23]. As with many nonlinear problems, we expect a close connection between amplitude and (here spatial) periodicity and we accordingly scale the horizontal coordinate to be proportional to amplitude. We also assume long time scales, but the factor by which we stretch the time is found essentially by trial and error. We let

$$\tilde{x} = \varepsilon x, \qquad \tilde{t} = \varepsilon^3 t.$$

Then we follow a deplorable notational trick that is widely used in fluid dynamics and drop the tildes. The reason is that the tildes make the equations cumbersome to read while the other letters we might have used for a rescaled time are preempted. The equations that we want to study now are

$$\theta_{zz} = \varepsilon^2 (\Psi_x - \theta_{xx} + \Psi_z \theta_x - \Psi_x \theta_z) + \varepsilon^3 \theta_t \tag{6}$$

$$\tau \Sigma_{zz} = \varepsilon^2 (\Psi_x - \tau \Sigma_{xx} + \Psi_z \Sigma_x - \Psi_x \Sigma_z) + \varepsilon^3 \Sigma_t \tag{7}$$

$$\begin{aligned} \Psi_{zzzz} = {} & R\theta_x - \tau S \Sigma_x + \varepsilon^2 [-2\Psi_{xxzz} + \frac{1}{\sigma}(\Psi_z \Psi_{zzx} - \Psi_x \Psi_{zzz})] \\ & + \frac{\varepsilon^3}{\sigma} \Psi_{tzz} + \varepsilon^4 [-\Psi_{xxxx} + \frac{1}{\sigma}(\Psi_z \Psi_{xxx} - \Psi_x \Psi_{xxz})] + \frac{\varepsilon^5}{\sigma} \Psi_{txx} . \end{aligned} \tag{8}$$

where the subscripts z,x,t signify partial differentiation with respect to the variables indicated.

Now we set out to find asymptotic solutions for small ε taking the other parameters, σ and τ, to be of order unity. This may not be the best choice for astrophysics, but it is often used by numerical experimenters and it provides a simple introduction.

EXPANSIONS

For small ϵ we consider expansions of the form,

$$\Psi = \Psi_0 + \epsilon\Psi_1 + \epsilon^2\Psi_2 + \dots ,$$
$$\theta = \theta_0 + \epsilon\theta_1 + \epsilon^2\theta_2 + \dots ,$$
$$\Sigma = \Sigma_0 + \epsilon\Sigma_1 + \epsilon^2\Sigma_2 + \dots .$$

The boundary conditions are

$$\theta_{nz}=0,\ \Sigma_{nz}=0,\ \Psi_n=0,\ \Psi_{nzz}=0 \quad @ \quad z=\pm\tfrac{1}{2}.$$

From (6)-(8) we obtain

$$\theta_{0zz} = 0, \quad \Sigma_{0zz} = 0 \tag{9a,b}$$

and

$$\Psi_{0zzzz} = (R_0 f - \tau S_0 g)_x . \tag{10}$$

The solutions are

$$\theta_0 = f(x,t), \quad \Sigma_0 = g(x,t) ,$$

where f and g are arbitrary functions to be determined and

$$\Psi_0 = (R_0 f - \tau S_0 g)_x \, P(z)$$

where

$$P(z) = \frac{1}{4!}\left(z^4 - \tfrac{3}{2} z^2 + \tfrac{5}{16}\right).$$

In exactly the same way we find that

$$\theta_1 = f_1(x,T), \quad \Sigma_1 = g_1(x,T)$$

where f_1 and g_1 are also arbitrary so far and

$$\Psi_1 = (R_0 f_1 - \tau S_0 g_1)_x P(z) .$$

Next we find that

$$\theta_{2zz} = \Psi_{0x} - \theta_{0xx} + \Psi_{0z}\theta_{0x} - \Psi_{0x}\theta_{0z}$$

$$\tau\Sigma_{2zz} = \Psi_{0x} - \tau\Sigma_{0xx} + \Psi_{0z}\Sigma_{0x} - \Psi_{0x}\Sigma_{0z}$$

which may be written as

$$\theta_{2zz} = (R_0 f - \tau S_0 g)_{xx} P - f_{xx} + f_x (R_0 f - \tau S_0 g)_x P' , \tag{11a}$$

$$\tau\Sigma_{2zz} = (R_0 f - \tau S_0 g)_{xx} P - \tau g_{xx} + g_x (R_0 f - \tau S_0 g)_x P' . \tag{11b}$$

Since (11a,b) are inhomogeneous forms of (9a,b), we know

that we have to remove the secular terms. If we integrate (11a,b) in z from $-\frac{1}{2}$ to $+\frac{1}{2}$ we find that the integrals on the left side vanish. Hence, so must the integrals on the right side, and the expression of that fact gives the solvability conditions for (11a,b). On noting that the integral of P is $(51)^{-1}$, we obtain the conditions

$$\mathbf{A}\begin{Vmatrix} f_{xx} \\ g_{xx} \end{Vmatrix} = 0; \quad \mathbf{A} = \begin{Vmatrix} R_0 - 51 & -\tau S_0 \\ R_0 & -\tau(S_0 + 51) \end{Vmatrix}. \tag{12}$$

For (12) to have nontrivial solutions we require that $\det \mathbf{A} = 0$, whence we find that

$$R_0 - S_0 = 51$$

This determines a critical value for the total Rayleigh number, R-S, which is just that found for the Bénard problem with fixed heat flux [9]. We also may note that the right and left null vectors of $\mathbf{A}$ are respectively

$$\mathbf{r} = \begin{Vmatrix} 1 \\ \tau^{-1} \end{Vmatrix}, \qquad \mathbf{l} = \| R_0, \ -S_0 \|.$$

We see that (12) is satisfied for

$$f = \tau g.$$

Now we may solve (11) and we obtain

$$\theta_2 = f_2(x,T) + H_2(z) f_{xx} + G_2(z) (f_x)^2$$

$$\tau^2 \Sigma_2 \quad \tau^2 g_2(x,t) + \tau H_2(z) f_{xx} + G_2(z) (f_x)^2,$$

where f_2 and g_2 are yet two more functions to be found and

$$G_2 = z^5 - \tfrac{5}{2} z^3 + \tfrac{25}{176} z$$

and

$$H_2 = \tfrac{1}{6} (z^6 - \tfrac{15}{4} z^4 + \tfrac{27}{176} z^2 - \tfrac{99}{8 \cdot 176}),$$

where an arbitrary constant has been chosen for later convenience.

Next from (8) we find, after some reductions, an equation for $\psi_{2_{zzzz}}$. This is easily solved and we obtain

$$\psi_2 = [R_0 f_2 - \tau S_0 g_2 + (R_2 - S_2) f]_x P(z) + P_2(z) f_{xxx} + Q_2(z)_x f_{xxx}$$

where

$$P_2 = \frac{(51)^2}{101} \left(z^{10} - \frac{45}{4} z^8 - \frac{483}{8} z^6 + \frac{18675}{64} z^4 - \frac{96345}{256} z^2 + \frac{9835}{128} \right).$$

Q_2 is an odd polynomial, which is all we need to know about it.

And so it goes. At ε^3 the time derivatives appear:

$$\theta_{3zz} = \psi_{1x} - \theta_{1xx} + \psi_{1z}\theta_{0x} + \psi_{0z}\theta_{1x} + \theta_{0t}$$

$$\tau\Sigma_{3zz} = \psi_{1x} - \tau\Sigma_{1xx} + \psi_{1z}\Sigma_{0x} + \psi_{0z}\Sigma_{1x} + \Sigma_{0t}$$

We integrate these two equations in z across the layer as we did in the previous orders. We obtain a pair of differential equations, or a matrix differential equation. We multiply this by **l**, the left null vector of **A**, and obtain the solvability condition

$$(R_0 - S_0/\tau)\, f_t = 0.$$

This means that either f_t or its coefficient must vanish. If the former is true, we have chosen the wrong time scale. Indeed that is so if we want certain kinds of solutions. For example, if $\tau=1$ or if both ζ_θ and ζ_Σ are positive, we expect steady convection and should scale the time accordingly. However, for the problem we are studying, we choose to let

$$R_0 - S_0/\tau = 0.$$

Then we have

$$R_0 = \left(\frac{1}{1-\tau}\right)5!, \qquad S_0 = \left(\frac{\tau}{1-\tau}\right)5! \tag{13a,b}$$

Having thus removed secular terms we may solve the two equations that result from the z-integration and we get

$$f_{1xx} - \tau g_{1xx} = -\frac{1-\tau}{\tau} f_t. \tag{14}$$

We are now near the end and may at last skip to order ε^4, which leads to the equations

$$\theta_{4zz} = \psi_{2x} - \theta_{2xx} + \psi_{0z}\theta_{2x} - \psi_{0x}\theta_{2z} + \psi_{1z}\theta_{1x} + \psi_{2z}\theta_{0x} + \theta_{1t} \tag{15a}$$

$$\tau\Sigma_{4zz} = \psi_{2x} - \tau\Sigma_{2xx} + \psi_{0z}\Sigma_{2x} - \psi_0\Sigma_{2z} + \psi_{1z}\Sigma_{1x} + \psi_{2z}\Sigma_{0x} + \Sigma_{1t}. \tag{15b}$$

Now we integrate (15a,b) over z and we find that

$$f_{1t} + \int_{-1/2}^{+1/2} [\psi_{2x} - \theta_{2xx} + \psi_{0z}\theta_{2x} - \psi_{0x}\theta_{2z}]\, dz = 0 \tag{16a}$$

and

$$g_{1t} - \int_{-1/2}^{+1/2} [-\psi_{2x} + \tau\Sigma_{2xx} - \psi_{0z}\Sigma_{2x} + \psi_{0x}\Sigma_{2z}]\, dz = 0. \tag{16b}$$

It is then straightforward to derive from (14) and (16a,b) an equation for f. The final step involves some integrations and we obtain then the evolution equation

$$f_{tt} - \mu\tau f_{xxxx} - \kappa\tau f_{xxxxxx} - \nu[(f_x)^3]_{xxx} = 0\,, \qquad (17)$$

where

$$\mu = (R_2 - S_2)/5!\,,$$

$$\int_{-1/2}^{+1/2} (P_2 - H_2)\,dz = .1967893\ldots \equiv \kappa,$$

and

$$(5!)^2 \int_{-1/2}^{1/2} P^2\,dz = 1.2301587\ldots \equiv \nu.$$

Equation (17) is a nonlinear wave equation whose properties we are attempting to understand. Here we sketch one of its approximate solutions that gives the flavor of the answers we seek.

BUOYANCY WAVES

When the amplitude of f is infinitesimal, the evolution equation may be linearized and it has a solution of the form

$$f = e^{\eta t}\cos(kx)\,.$$

This gives us

$$\eta^2 = \tau k^4(\mu - \kappa k^2)\,.$$

So we have instability whenever

$$\mu \geq \mu_0 \equiv \kappa k^2\,.$$

If the situation is only slightly unstable we can once again make an amplitude expansion. Let

$$\mu = \mu_0 + \frac{\lambda}{\tau}\delta^2$$

where $\delta^2 << 1$ and λ is an arbitrary parameter; δ is analogous to ε in the previous development and λ is analogous to $R_2 - S_2$. Hence we scale the amplitude with δ and set

$$f = \delta F.$$

We also define a slow time

$$s = \delta t.$$

The evolution equation becomes

$$\kappa\tau(F_{xxxxxx} + k^2 F_{xxxx}) = \delta^2\{F_{ss} - \lambda F_{xxxx} - \nu[(F_x)^3]_{xxx}\}.$$

We expand again:

$$F = F_0 + \delta F_1 + \delta^2 F_2 + \ldots .$$

In leading order we find the linear problem

$$F_{0xxxxxx} + k^2 F_{0xxxx} = 0 ,$$

with the solution

$$F_0 = X(s)\cos(kx) + Y(s)\sin(kx) ,$$

where X and Y are arbitrary functions. Then in the next order,

$$\kappa\tau(F_{1xxxxxx} + b^2 F_{1xxxx})$$
$$= (\ddot{X} - \lambda k^4 X)\cos(kx) + (\ddot{Y} - \lambda k^4 Y)\sin(kx)$$
$$- \nu k^3[X\sin(kx9 - Y\cos(kx)]_{xxx} .$$

We multiply by sin(kx) and integrate from 0 to $2\pi/k$; then we multiply by cos(kx) and integrate. This leads to coupled equations for X and Y. Rather than write these directly we prefer to use as variables A and ϕ where

$$X = A\cos\phi, \quad Y = A\sin\phi.$$

Then

$$F = A\cos(kx + \phi).$$

The equations are

$$\ddot{A} - A\dot{\phi}^2 - \lambda k^4 A - \tfrac{3}{4}\nu k^5 A^3 = 0 .$$

and

$$A\ddot{\phi}^2 + 2\dot{A}\dot{\phi} \quad 0 .$$

We find that

$$\dot{\phi} = b/A^2$$

where b is arbitrary and we get the amplitude equation

$$\ddot{A} - b^2/A^3 - \lambda k^4 A - \tfrac{3}{4}\nu k^5 A^3 = 0 . \qquad (18)$$

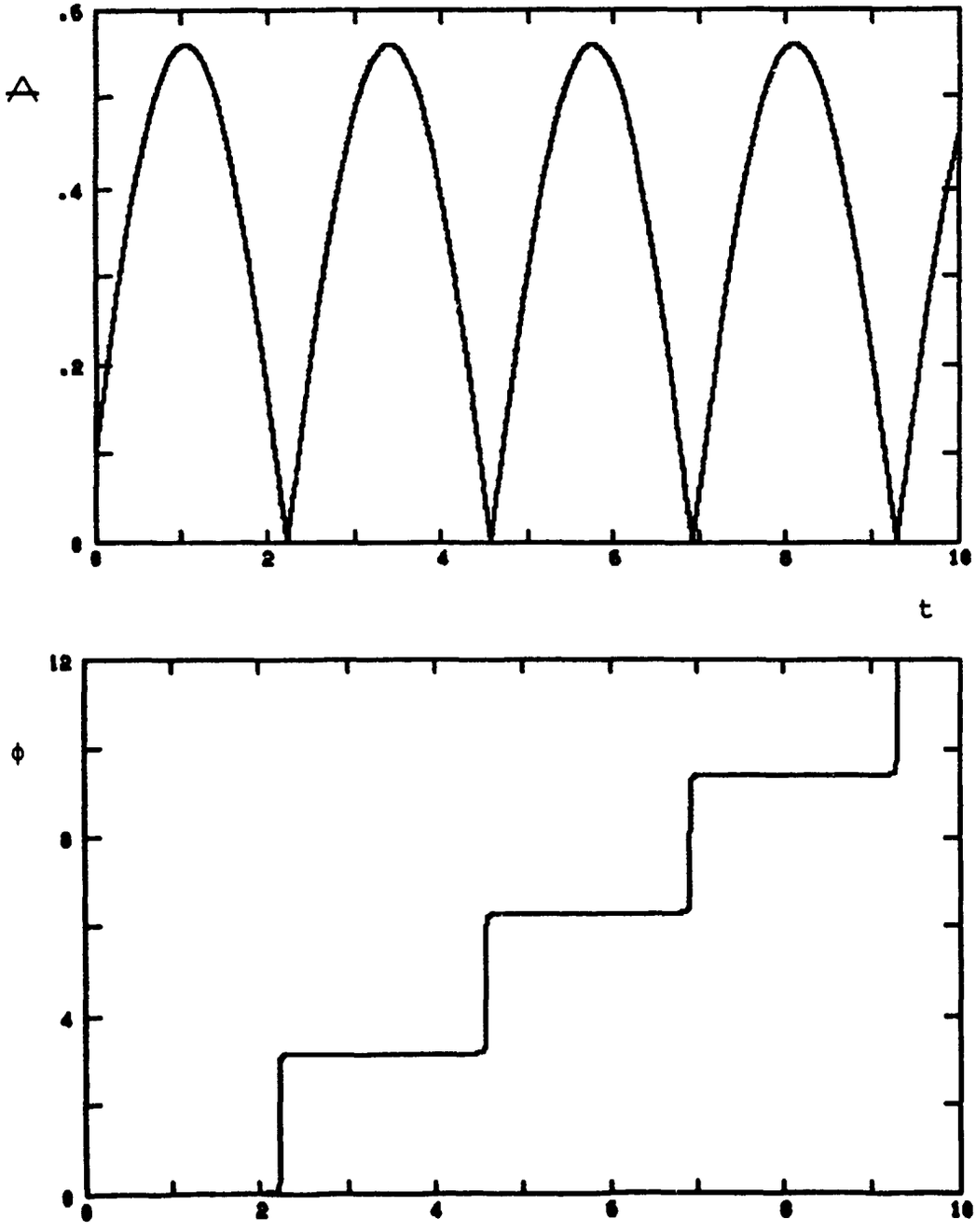
This has the integral

$$\tfrac{1}{2}\dot{A}^2 + V(A) = E$$

where

$$V = \tfrac{1}{2}\left(b^2/A^2 - \lambda k^4 A^2 - \tfrac{3}{8}\nu k^5 A^4\right)$$

and E is a constant. Solutions may be expressed in elliptic functions, but it is instructive simply to look at plots of the amplitude and phase, here for b = .001, $\lambda = -2$ and k = 1.



THE CAVEATS

We have suggested that a certain kind of variability is caused by instabilities in thin layers on large horizontal scales with long periods. This opens the way for an analysis such as we have just sketched. The calculations may be elaborate, but they are feasible. Their astrophysical interest lies in the relative ease with which they may be extended to allow for nonBoussinesq effects [25,26], when those are not too pronounced. Thus one can treat compressibility, κ-mechanism, and so on, and the extension to more general boundary conditions has been studied [27]. In other words, a number of features of nonlinear nonradial pulsation can be studied along the lines we have outlined here. It only requires finding the right instability.

However, the main outcome so far is only qualitative because (18) has an infinite number of possible solutions and we have not given a method for selecting one from among them. The removal of this degeneracy requires the introduction of higher order information and this may be effected by procedures that we shall go into elsewhere. Nevertheless, the cyclic character of the solutions is a correct, if particular, asymptotic consequence of the equations.

Yet equation (18) is too tame and it gives only periodic solutions. Nor do the higher order corrections remove this failing. But we do have a situation with two competing instabilities such as the models mentioned at the beginning did. If they give chaotic solutions [5,28] why do we not find them from (18)? The formal difference is simply that (18) is second order whereas the model equations are third order. If we analyze the latter in the neighborhood of the onset of the instabilities, we may reduce them to second order equations. Chaos in the models arises only for highly unstable conditions and was discovered by numerical means. If we want to study strongly nonlinear conditions in the problems discussed here, we have to solve nonlinear partial differential equations, and that is a far more difficult task than solving the model equations, which are ordinary differential equations. Perhaps when the full numerical solutions are found, there will appear just the rich structure we see in the solar cycle; that is certainly one thing that should be attempted. But there is another way to enrich the time dependence of our model that we believe is relevant to the solar case and must be included in any event.

If we consider a case with three competing instabilities, the procedures described here lead to a third order amplitude equation near a critical point at which all three instabilities begin at once. We have been studying these questions with colleagues in Nice and that work [29] will provide some concrete examples of what we mean. For our present purposes we need only the simple extension of the idea of competing instabilities to the case of three instablities. It turns out that in each of the examples of

variable variability cited here, there seem to be three competing instabilities (at least potentially) in the models that are being proposed. Consider the case of solar variability and the picture that it is driven from the base of the convective zone [13].

In the process of solar spin down [30], it has been suggested, when the hydromagnetic torques of the solar wind brake the convective zone, it in turn pumps a secondary flow into the subconvective layers [31]. The convective pumping process is resisted by the stable layers, hence it can penetrate only into a shallow layer below the convective zone [30,32]. This layer initially supports the rotational difference between the convective zone and the radiative interior, but it ultimately loses stability. The further developments are not fully understood, but one of the plausible possibilities is that the resulting motions maintain the layer in conditions near to marginal instability. This is the source of the third competing instability that we believe must be included in the description of the solar activity cycle. For schematic versions of this problem, if the geometry is right, one gets third order equations for the amplitude of the motion. Therein, we suggest, lies the cause of some of the chaos of the solar cycle.

This hint of further developments only underscores how incomplete is the picture we have presented here. But at least we have been able to see one direction to go in which to find the source of variable variability. The temporal behaviors that are emerging at this stage of the work have some of the right kind of qualitative behavior and the mechanism of competing instabilities seems to provide a possible basis for understanding the examples of stellar variability that we have mentioned here.

REFERENCES

1. Eddy, J. A.: Historical and Arboreal Evidence for a Changing Sun, The New Solar Physics, ed. J. A. Eddy, 1978, Westview Press, Boulder, pp. 11-34.

2. Robinson, E. L.: The Observational Properties of the ZZ Ceti Stars, IAU Colloquium No. 53, White Dwarfs and Variable Degenerate Stars, ed. H. M. Van Horn and V. Weicheman, 1979, Univ. of Rochester, Rochester, pp. 343-357.

3. Scherrer, P. H.; Wilcox, J. M.; Kotov, V. A.; Severny, A. B.; and Tsap, T. T.: Observations of Solar Oscillations with Periods of 160, Nature, vol. 277, 1979, pp. 635-637.

4. Marsden, J. E. and McCracken, M.: The Hopf Bifurcation and its Applications, 1976, Springer-Verlag, New York.

5. Moore, D. W. and Spiegel, E. A.: A Thermally Excited Non-Linear Oscillator, Astrophys. J., vol. 143, 1966, pp. 871-887.

6. Guckenheimer, J.: Multiple Bifurcation Problems of Co-demension Two, 1981, Preprint, University of Santa Cruz.

7. Baker, N. H.; Moore, D. W.; and Spiegel, E. A.: Nonlinear Oscillations in the One-Zone Model for Stellar Pulsation, Astron. J., vol. 71, 1966, p. 844(A).

8. Chandrasekhar, S.: Hydrodynamic and Hydromagnetic Stability, 1961, Clarendon Press, Oxford.

9. Hurle, D. T. G.; Jakeman and Pike, E. R.: On the Solution of the Bénard problem with boundaries of finite conductivity, Proc. Roy. Soc., A. vol. 296, 1967, p. 469-475.

10. Dziembowski, W. and Koester, D.: Excitation of Gravity Modes in White Dwarfs with Chemically Stratified Envelopes,1980, Preprint.

11. Dolez, N.: Thesis, 1981 Observatoire de Paris, Meudon, in preparation.

12. Parker, E. N.: Cosmical Magnetic Fields: Their Origin and Activity, 1979, Clarendon Press, Oxford.

13. Spiegel, E. A. and Weiss, N. O.: Magnetic activity and variations in solar luminosity, Nature, vol. 287, 1980, pp. 616-617.

14. Moffatt, H. K.: The Generation of Magnetic Fields in Electrically Conducting Fluids, 1978, Cambridge University Press.

15. Gough, D. O.: On the seat of the Solar Cycle, Workshop on solar constant variations, NASA CP-2191, 1981.

16. Witham, G. B.: Linear and Nonlinear Waves, 1978, John Wiley and Sons, New York.

17. Eckart, C.: Hydrodynamics of Oceans and Atmospheres, 1960, Pergamon, London.

18. Dilke, F. N. W. and Gough, D. O.: The Solar Spoon, Nature, vol. 200, 1972, pp. 262-264.

19. Spiegel, E. A. and Weiss, N. O.: Magnetic Buoyancy and the Boussinesq Approximation, 1981, Submited to J. Geophys. Astrophys. Fluid Dyn.

20. Spiegel, E. A.: Semiconvection, Comm. Astrophys and Space Phys., vol 1, 1969, pp. 57-64.

21. Turner, J. S.: Buoyancy Effects in Fluids, 1973, Cambridge University Press.

22. Childress, S. and Spiegel, E. A.: Pattern Formation in a Suspension of Microorganisms: Nonlinear Aspects, 1978, submitted to J. Fluid Mech.

23. Chapman, G. J. and Proctor, M. E.: J. Fluid Mech. 19 vol. 101, 1980, pp. 759-782.

24. Huppert, H. E. and Moore, D. R.: Nonlinear double-diffusive convection, J. Fluid Mech. vol. 78, 1976, pp. 821-854.

25. Depassier, M. C. and Spiegel, E. A.: The Large-Scale Structure of Compressible Convection, Astron. J., vol. 86, 1981, pp. 196-512.

26. Depassier, M. C. and Spiegel, E. A.: Convection with heat flux prescribed on the boundaries of the system, I. The effect of temperature dependence of material properties, submited to J. Geophys. and Astrophys. Fluid Dyn., 1981.

27. Poyet, J.-P.: Thesis, 1980, Department of Astronomy, Columbia University, New York.

28. Marzec, C. J. and Spiegel, E. A.: Ordinary differential equations with strange attractors, SIAM J. App. Math, vol. 38, 1980, pp. 387-421.

29. Howard, L. N.; Moore, D. W.; and Spiegel, E. A.: The Solar Spin-Down Problem, Nature, vol. 214, 1967, pp. 1297-1300.

30. Bretherton, F. P. and Spiegel, E. A.: The Effect of the Convection Zone on Solar Spin-Down, Astrophys. J., vol. 153, pp. L153-L158.

31. Spiegel, E. A.: A History of Solar Rotation, Physics of the Solar System, ed. S. I. Rasool, 1972, NASA SP-300, pp. 61 ff.

WORKSHOP ON VARIATIONS OF THE SOLAR CONSTANT

ATTENDEES

Arthur Aikin
Code 964, Bldg. 21, Room 262
Goddard Space Flight Center
Greenbelt, MD 20771

Dr. S. P. S. Anand
Applied Research Corporation
8401 Corporate Drive
Landover, MD 20785

Dr. Ronald Angione
San Diego State University
San Diego, CA 92115

Albert Arking
Code 915, Bldg. 22, Room 122
Goddard Space Flight Center
Greenbelt, MD 20771

J. Bahcall
Princeton Institute of
Advanced Studies
Olden Lan, Bldg. E
Princeton, NJ 08540

Siegfried Bauer
Code 600, Bldg. 26, Room 200
Goddard Space Flight Center
Greenbelt, MD 20771

P. R. Bell
Oak Ridge Institute for Energy
Analysis - Assoc. Univ.
P.O. Box 117
Oak Ridge, TN 37830

W. Blitzstein
University of Pennsylvania
Department of Astronomy
Philadelphia, PA

James Burcham
Code 961, Bldg. 21, Room 259
Goddard Space Flight Center
Greenbelt, MD 20771

K. Chan
Code 961
Goddard Space Flight Center
Greenbelt, MD 20771

D. Dearborn
University of Arizona
Tucson, AZ 85721

R. Donnelly
NOAA - R43
U.S. Department of Commerce
Environmental Research Laboratory
Boulder, CO 80303

Maurice Dubin
Code 963, Bldg. 21, Room 272
Goddard Space Flight Center
Greenbelt, MD 20771

C. Duncan
Code 942, Bldg. 22, Room 386
Goddard Space Flight Center
Greenbelt, MD 20771

D. Dunham
System/Science Division
Computer Science Corporation
8728 Colesville Road
Silver Spring, MD 20910

L. Dunkelman
University of Arizona
Tucson, AZ 85721

Mary Durfman
Science Magazine
1515 Massachusetts Avenue, NW
Washington, DC 20005

T. Duvall
Kitt Peak National Observatory
P.O. Box 26732
Tucson, AZ

Igor Eberstein
Code 964, Bldg. 21, Room 261
Goddard Space Flight Center
Greenbelt, MD 20771

John Eddy
National Center for Atmospheric Research
Boulder, CO 80302

Andrew Endal
Department of Physics and Astronomy
Louisiana State University
Baton Rouge, LA 70803

A. Fiala
U.S. Naval Observatory
Washington, DC 20390

P. Foukal
AER, Inc.
872 Massachusetts Avenue
Cambridge, MA 02139

C. Frohlich
Physico-Meteorological Observatory
World Radio Team Center
Davos, Switzerland

Bertrand Gagnon
Code 960, Bldg. 21, Room 208
Goddard Space Flight Center
Greenbelt, MD 20771

J. Geist
National Bureau of Standards
Gaithersburg, MD

H. Glaser
NASA Headquarters
Goddard Space Flight Center
Greenbelt, MD 20771

Richard Goldberg
Code 961, Bldg. 21, Room 277B
Goddard Space Flight Center
Greenbelt, MD 20771

Dr. Douglas Gough
Institute of Astronomy
Cambridge University
Madingley Road
Cambridge CB30HA
England

R. E. Hartle
Code 961, Bldg. 21, Room 232
Goddard Space Flight Center
Greenbelt, MD 20771

J. Harvey
Kitt Peak National Observatory
P.O. Box 26732
Tucson, AZ 85726

Dr. D. Heath
Code 963, Bldg. 21, Room 268
Goddard Space Flight Center
Greenbelt, MD 20771

J. R. Hickey
Eppley Laboratories
Newport, RI 02840

H. Hill
Physics Department, Bldg. 81
University of Arizona
Tucson, AZ 85721

Dr. Robert Kandel
Service d'Aeronomie du C.N.R.S.
91370 Verrieres-leBuisson
France, BP. No. 3

J. Kendall
California Institute of Technology
Pasadena, CA 91125

Dr. R. Koch
Department of Astronomy
University of Pennsylvania
Philadephia, PA

B. LaBonte
Mt. Wilson
Las Campanas Observatory
813 Santa Barbara Street
Pasadena, CA 91101

Dr. William Livingston
Kitt Peak National Observatory
P.O. Box 26732
Tucson, AZ 85726

Dr. G. Wesley Lockwood
Lowell Observatory
Box 1269
Flagstaff, AZ 86002

Stephen Maran
Code 680, Bldg. 21, Room 138
Goddard Space Flight Center
Greenbelt, MD 20771

Hans Mayr
Code 961, Bldg. 21, Room 273
Goddard Space Flight Center
Greenbelt, MD 20771

Ann Mecherikunnel
Code 942, Bldg. 22, Room 264
Goddard Space Flight Center
Greenbelt, MD 20771

L. H. Meredith
Code 930, Bldg. 28, Room W231D
Goddard Space Flight Center
Greenbelt, MD 20771

J. M. Mitchell
NOAA/Environmental Data Services
Room 608
Gramax Building
Silver Spring, MD 20910

Dr. S. Modali
6404 Fairborn Terrace
New Carrolltown, MD 20784

D. Murcray
University of Denver
Denver, CO 80219

Dr. Gerald R. North
Code 915, Bldg. 22, Room G10
Goddard Space Flight Center
Greenbelt, MD 20771

John O'Keefe
Code 681, Bldg. 21, Room G34
Goddard Space Flight Center
Greenbelt, MD 20771

Ludwig Oster
Code 961, Bldg. 21, Room 255
Goddard Space Flight Center
Greenbelt, MD 20771

Frank D. Paluconi
Code EBT-8
NASA Headquarters
400 Maryland Avenue, SW
Washington, DC

Hongwoo Park
Systems & Applied Sciences Corp.
6811 Kenilworth Avenue
Riverdale, MD 20737

S. I. Rasool
IBM Scientific Center
36 Ave. Raymond Poincare
Paris 75116
France

Edward Rhodes
Department of Astronomy
UCLA
Los Angeles, CA 90024

R. Rood
Astronomy Department
University of Virginia
Box 3818, University Station
Charlottesville, VA 22903

Sr. I. W. Roxburg
Department of Applied Mathematics
University of London/Queen Mary College
Mile End Road
London E1-4NS, England

Sabatino Sofia
Code 961, Bldg. 21, Room 255
Goddard Space Flight Center
Greenbelt, MD 20771

Dr. K. Schatten
Code 961, Bldg. 21, Room 277B
Goddard Space Flight Center
Greenbelt, MD 20771

Dr. Barry Schlesinger
Systems & Applied Sciences Corp.
6811 Kenilworth Avenue
Riverdale, MD 20737

Myron Smith
Sacramento Peak Observatory
Sunspot, NM 88349

Clayton Smith
U.S. Naval Observatory
Washington, DC

E. Spiegel
Astronomy Department
Columbia University
New York, NY 10027

Richard Stolarski
Code 964, Bldg. 21, Room 228
Goddard Space Flight Center
Greenbelt, MD 20771

Dr. M. Stuiver
Quaternary Research Center
AK 20 - University of Washington
Seattle, WA 98195

Dr. A. Sweigart
Code 681, Bldg. 21, Room G28
Goddard Space Flight Center
Greenbelt, MD 20771

J. Thomas
Department of Aerospace Engineering
University of Rochester
Rochester, NY

G. Thuillier
Service d'Aeronomie de CNRS
91370 Verrieres-le-Buisson
France, BP No. 3

L. Twigg
Louisiana State University
Baton Rouge, LA 70803

Jack Van Zant
Code 960, Bldg. 21, Room 208
Goddard Space Flight Center
Greenbelt, MD 20771

D. Wentzel
Astronomy Branch
University of Maryland
College Park, MD 20742

Dr. Robert Willson
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91125

Dr. Charles L. Wolff
Code 961, Bldg. 21, Room 277A
Goddard Space Flight Center
Greenbelt, MD 20771